

USING MATRIX TRAINING TO ESTABLISH THE ALPHABETIC PRINCIPLE,
AND GENERALIZATION TO READING, IN TYPICALLY DEVELOPING
STRUGGLING READERS

By

Copyright 2016

Carol Cummings

Submitted to the graduate degree program in Applied Behavioral Science and the
Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for
the degree of Master of Arts.

Co-Chairperson Pamela L. Neidert, Ph.D.

Co-Chairperson Kathryn J. Saunders, Ph.D.

Vincent T. Francisco, Ph.D.

Date Defended: November 30, 2016

The Thesis Committee for Carol Cummings

certifies that this is the approved version of the following thesis:

USING MATRIX TRAINING TO ESTABLISH THE ALPHABETIC PRINCIPLE,
AND GENERALIZATION TO READING, IN TYPICALLY DEVELOPING
STRUGGLING READERS

Co-Chairperson Pamela L. Neidert, Ph.D.

Co-Chairperson Kathryn J. Saunders, Ph.D.

Date Approved: November 30, 2016

Abstract

Previous studies, primarily with adults with intellectual disabilities, demonstrated recombinative generalization of onset and rime units using matrix training. Study 1 extends that work to typically developing children with reading difficulties. Three boys, aged 4 to 8 years, participated. Word subsets containing all combinations of two onsets and two rimes (e.g. bed, bag, ked, kag) were taught using a computerized, matching-to-sample (MTS) task. Participants learned to select printed words that corresponded to spoken words, from a choice pool containing all words in a subset. Study 2 taught abstraction of phonemes *within the rime*. Each subset contained all combinations of two vowels and two codas (e.g., bed, beg, bad, bag). In both studies, participants showed generalization from MTS to reading words, and to MTS with untaught subsets. Generalization to untaught words demonstrates the “alphabetic principle”—the concept that the same sound in different words is represented by the same letter.

Tables of Contents

Abstract	iii
Table of Contents	iv
List of Illustrative Materials	v
Appendix Index	ix
Introduction	1
Study 1 Method	11
Study 1 Results	19
Study 1 Discussion	24
Study 2 Method & Results	28
Study 2 Discussion	35
General Discussion	38
References	41
Figure Captions	45
Tables	49
Figures	52
Appendices	67

List of Illustrative Materials

Table 1. Participant demographics.	49
Table 2. The word matrices used in Study 1.	50
Table 3. The word matrices used in Study 2.	51
<i>Figure 1.</i> The parts of a syllable, <i>bat</i> .	52
<i>Figure 2.</i> The six phases of Study 1, shown in order from top to bottom.	53
<i>Figure 3.</i> The trial layout for the spoken-to-printed word MTS task (top), prompted MTS trials (see Appendix), and the printed- word identity matching (bottom) are shown as progressing from left to right.	54
<i>Figure 4.</i> The percentage of correct responses made by MH (left) and KR (right) during the MTS Comprehensive Test before and after training of the - ag/-ed matrix. The rimes at the top left of each panel indicate the matrix represented. The black line represents –ag/-ed training.	55
<i>Figure 5.</i> The percentage of correct responses made by MH (left), TS (middle), and KR (right) during the Reading Comprehensive Test are shown. The vertical black line represents -ag/-ed Matrix Training.	56
<i>Figure 6.</i> The percentage of correct responses made by MH (top), TS (middle), and KR (bottom) on the baseline MTS Comprehensive Test (black columns) and the MTS Generalization Tests (crossed columns) given during Matrix Training for each subset in the -ag/-ed matrix. Subsets are shown in order of training from left to right on the x-axis.	57
*Accuracy on the RL subset for TS and KR are based on rimes.	
<i>Figure 7.</i> The percentage of correct responses made by MH (top), TS (middle),	58

and KR (bottom) on the baseline Reading Comprehensive Test (black columns) and the Reading Generalization Tests (crossed columns) given during Matrix Training for each subset in the -ag/-ed matrix. Subsets are shown in order of training from left to right on the x-axis.

*Accuracy on the RL subset for TS and KR are based on rimes.

Figure 8. The number of whole word responses (out of 12 opportunities) that were preceded by subsyllable sounds (y-axis) during the Reading Comprehensive Test given following mastery of the -ag/-ed matrix. TS is shown on the top panel, and KR on the bottom. The rime units of each matrix are presented across the x-axis. Each column shows the number of whole word responses that were preceded by phonemes (grey) and onset/rime sounds (checkered). This figure does not reflect accuracy of responses. 59

Figure 9. The number of trials (y-axis) needed for MH (top), TS (middle), and KR (bottom) to reach mastery criteria on MTS (black columns) and Reading (crossed columns) for each subset (x-axis). Subsets are indicated by their onsets and are shown in order of training from left to right. *Accuracy on the generalization tests for the RL subset for TS and KR were based on rimes only. 60

Figure 10. Accuracy on the Reading Comprehensive Test before and after training for MH (left) and KR (right). Matrices are indicated by the vowels at the top left of each panel. Training is indicated by a black vertical line. 61

<i>Figure 11.</i> Accuracy on the MTS Comprehensive Test before and after training for MH (left) and KR (right). Matrices are indicated by the vowels at the top left of each panel. Training is indicated by a black vertical line.	62
<i>Figure 12.</i> Accuracy on the baseline MTS Comprehensive Test (black columns) and the MTS Generalization Test (crossed columns) for each subset, indicated by the onset across the x-axis, for MH (top) and KR (bottom). The arrow in MH's graph represents when the Reading Probe was given. *There was not an MTS Generalization Test for the N Subset.	63
<i>Figure 13.</i> The percentage of correct responses made by MH (top) and KR (bottom) on the baseline Reading Comprehensive Test (black columns) and the Reading Generalization Tests (crossed columns) given during Matrix Training for each subset trained in Study 2. Subsets are shown in order of training from left to right on the x-axis. The arrow in MH's graph represents when the Reading Probe was given (see Appendix). *Due to time constraints, KR did not receive MTS training on the G subset, therefore his Reading Generalization Test score for that subset is not shown.	64
<i>Figure 14.</i> The number of KR's whole word responses (out of six opportunities) that were preceded by subsyllable sounds during the Reading Comprehensive Test given before (top) and after (bottom) training. The rime units of each matrix are presented across the x-axis. For each rime there is a column showing the number of responses to words ending in that rime that were preceded by phonemes (grey), onset/rime	65

sounds (checkered). This figure does not reflect accuracy of responses.

Figure 15. The number of trials (y-axis) needed for MH (top) and KR (bottom) to reach mastery criteria on MTS (black columns) and Reading (crossed columns) for each subset (x-axis). Subsets are indicated by their onsets and are shown in order of training from left to right. *G subset was not trained for reading due to time constraints. 66

Figure 16. A scoresheet, with added column for “Nature of Error,” used to show the possible spoken responses a participant can give and how the experimenter should record them during the session. 78

Appendix Index

A) Study 1 MTS Training Procedures	67
B) Study 1 Reading Training Procedures	70
C) Study 2 MTS Training Procedures	72
D) Study 2 Reading Training Procedures	74
E) Scoring Procedures	77

Using Matrix Training To Establish The Alphabetic Principle, and Generalization To Reading, In Typically Developing Struggling Readers

Reading is one of the most important skills to master, yet many students are struggling to learn how to read. The National Assessment of Education Progress (NAEP) is the only national assessment that periodically reports students' performance in mathematics and reading level (NAEP, 2015). The assessment offers three levels of achievement on each subject: basic, proficient, and advanced. Fourth grade is the first grade at which this assessment is given. According to the NAEP, fourth graders performing at or above the basic level for reading can locate relevant information in text, and therefore can use their reading skills to acquire more knowledge. In 2015, the NAEP reported that 31% of fourth graders read below the basic level.

Because it has been shown that those who are behind their peers in fourth grade are likely to remain behind (Stanovich, 1986), it is important to establish and monitor mastery of component skills for reading before then. An efficient, evidence-based instructional program designed to teach the necessary component skills for reading may be useful in preventing students from falling behind their peers. This study investigates the effectiveness of an instructional program designed to teach the alphabetic principle, discussed later, to students at the early stages of reading instruction.

Reading is a complex behavior that requires mastery of several skills, with a critical goal being the ability to produce the sound of a word upon seeing it. Adams (1990) explained, "Unless the processes involved in individual word recognition operate properly, nothing else in the system can either" (p. 3). However, any language contains too many words to learn by sight (i.e., memorizing the sound of each whole word).

Teaching whole words by sight limits students to reading *only* those words that have been taught. It is more logical to use procedures designed to establish the skills needed to read words that have not been taught, referred to as decoding, or Word Attack (Adams, 1990; NAEP, 2015; Schuele & Boudreau, 2008). These decoding or Word Attack skills require the alphabetic principle and phonemic abstraction (Snow, Burns, & Griffin, 1998).

The Alphabetic Principle and Phonemic Abstraction

Byrne & Fielding-Barnsley (1998) defined the alphabetic principle as “useable knowledge of the fact that phonemes can be represented by letters, such that whenever a particular phoneme occurs in a word, and in whatever position, it can be represented by the same letter” (p. 313). As Figure 1 shows, syllables are broken into onsets (first letter) and rimes (vowel and remaining consonants), which can be further broken into phonemes. Phonemes are the smallest spoken speech sounds. The alphabetic principle requires phonemic abstraction (PA), referred to as phonemic awareness in the reading literature. When a student is capable of isolating, or segmenting, one phoneme and recognizing it across words that have not been directly taught, this student is demonstrating PA.

The National Reading Panel (NRP) defines PA as “the ability to focus on and manipulate phonemes in spoken words” (2000, p. 2-1). An example is recognizing that the /a/ sound in *bat* is the same as in *rad*, or that the /b/ sound in *bed* is the same in *bun*. Evidence gathered by reading researchers shows that PA plays a critical role in learning to read (Byrne, 1994; National Reading Panel [NRP], 2000; NAEP, 2015).

Catania defines abstraction as “discrimination based on a single stimulus property independent of other properties: thus generalization among all stimuli with that property” (1998, p. 378). Therefore, one can test for PA by testing for generalization to untaught, spoken words. For example, you would first teach a student that *bed*, *burger*, and *bent* all start with the same sound, /b/, and that *sit*, *super*, and *sack* all start with the sound /s/. In a test for generalization, you would ask the student, “Which word starts with the same sound as *bed*, is it *sun* or *basket*?” Generalization is said to occur if the student’s response is “*basket*.” This shows that the student is able to isolate and discriminate between the sounds /s/ and /b/ within words that have not been taught.

Byrne & Fielding-Barnsley (1989) conducted a series of experiments to determine conditions needed to acquire the alphabetic principle. The second experiment in the series was conducted to determine effects of directly teaching two PA skills, segmentation and sound categorization, on acquisition of the alphabetic principle. The tasks described below were used with twelve typically developing children. The mean age of the children was 54.5 months; they could name an average of 8.3 letters, and accurately produce an average of 1.3 letter sounds. The word pairs used in training and generalization testing each consisted of one word starting with *m*, and one starting with *s*. The rimes of both words in each pair were the same (e.g., *mat* and *sat*, *mop* and *sop*, *mow* and *sow*). The generalization test, given after each skill was taught, was used to determine whether the alphabetic principle was acquired.

The first skill taught, segmentation, requires PA. On each trial, participants repeated a spoken sample, “*mat*,” with the onset said separately from the rime, “*m...at*.” Participants had to respond correctly on three consecutive trials before moving on to the

next word. Ten words were trained: *mat*, *sat*, *mum*, *sum*, *met*, *set*, *mad*, *sad*, *mow*, and *sow*. Of the twelve participants, eight learned to segment the words, and four did not.

Next, participants were taught to read two words, *mat* and *sat*. Participants could move on after responding correctly on six consecutive trials. All twelve participants learned to read the two printed words.

The generalization test was first given following reading training with *mat* and *sat*. The test included all of the words used in segmentation training, except *mat* and *sat*. The experimenter presented a printed word and asked, “Does this say *mum* or *sum*?” while showing two corresponding pictures, a woman (*mum*) and a plus sign (*sum*), from which the participant could choose. Each word was presented in this way once. None of the participants met the criterion of seven of eight correct trials, thus showing they had not acquired the alphabetic principle.

Following the generalization test, participants began sound-categorization training. They were presented several spoken words beginning with either *m* or *s* (e.g., *milk*, *mother*, *Sally*, *sandwich*), and told what sound each word started with. The experimenter then asked, “What word starts with the same sound as *mat*, is it *mum* or *sum*?” The word “*sat*” was then presented in the same way, again with “*mum*” and “*sum*” as the choices. The correct answer was provided if the participant was incorrect. This continued until the four word pairs from the generalization test had been presented in the same format, and the participant responded correctly on six of the eight questions. Success in this phonemic abstraction task would mean that the participant can abstract the sounds /m/ and /s/ across words. Only five of the twelve participants learned this task.

The generalization test was then given to all twelve again. Results of the test showed none of the participants had acquired the alphabetic principle.

The last task was letter-sound training. The participants were shown the printed letters *m* and *s*, and told their corresponding sounds. This was repeated several times. They were then required to say the corresponding sounds when shown *m* and *s* individually until they made six consecutive correct responses. All twelve participants met criterion before receiving the generalization test again.

Six of the twelve participants met criterion on the last generalization test. Of the six successful participants, all had met criterion on at least one of the phonemic-abstraction tasks. Four reached criterion on both segmentation and sound categorization training, one reached criterion on segmentation, and one reached criterion on sound categorization training. Those participants who did not meet criterion on at least one of the PA tasks failed to show generalization on the last test for the alphabetic principle.

The authors concluded that both PA and letter-sound training, in combination, promote acquisition of the alphabetic principle. The current study's teaching and testing procedures were developed from the procedures used by Byrne & Fielding-Barnsley (1989) to test for the alphabetic principle.

Characteristics of Effective Instructional Programming

An evidence-based instructional program may be used to teach component reading skills to students. In developing such a program, we consider several key features discussed by Holland (1960) and Keller (1968). Holland and Keller summarized research on effective teaching procedures in which both have discussed the importance of providing many opportunities for responses to be made, immediate reinforcement of

those responses, gradual progression when establishing complex skills, and individualization. The computerized program used in this study incorporated several of these key features. For example, the computerized program allowed for a large number of trials in a short amount of time, roughly 5 trials per minute. The touch screen feature allowed for a clear response that is reinforced immediately. And the structure of the matrix training required that each participant master a subset before moving on.

In addition to important key features of the instructional program, how best to program for phonemic abstraction was also considered. Skinner (1957) explained that when a verbal response is emitted and reinforced in the presence of a stimulus, that stimulus controls that verbal response. To promote abstraction, verbal responses should be differentially reinforced based on single a characteristic of a stimulus (Skinner, 1957). To do this, one must present several examples of the characteristic, as well as several non-examples. This is important when programing for abstraction because it isolates the relevant characteristic that you wish for the student to abstract.

For instance, you teach a student to say “*triangle*” in the presence of a blue triangle, so that their verbal response will come under the control of the presented shape. When shown a pink triangle, however, the student does not say “*triangle*,” but does say “*triangle*” when they see something blue. The color, blue, controls the student’s verbal response, “*triangle*.” To change this, the student must be presented with triangles of several colors, and other shapes of various colors. Saying “*triangle*” should be reinforced only when presented with triangles, regardless of color. This provides several examples (triangles) and non-examples (squares, rectangles, circles, etc.) of the relevant stimulus

characteristic, shape. It also provides examples of the non-relevant stimulus characteristic, color.

Matrix Training

Matrix training is a form of instruction that promotes abstraction. It is used to produce recombinative generalization, defined as the combining of at least two known stimuli in a novel way (Goldstein, 1983a). A matrix is composed of all combinations of at least two stimulus components. Once several combinations have been learned, novel combinations can be acquired without additional training. Some previous studies using matrix training have focused on labeling (e.g. Whitehurst, 1971), instruction following (e.g. Striefel & Wetherby, 1973; Goldstein, Angelo, & Mousetis, 1987), and reading (e.g. Hanna et al., 2011). These studies typically involve the recombination of discrete whole words (e.g. Goldstein, 1983; Goldstein & Mousetis, 1989; Striefel et al. 1976, 1978).

Mineo & Goldstein (1990) showed recombinative generalization of whole words with four children with language delays using matrix training. They used a 6x8 matrix of “action words” and “object words” to teach the children to follow action-object instructions. The authors first receptively trained all combinations of two action words (*point to* and *push*) and four object words (*spoon*, *car*, *cup*, and *shoe*). For example, the participants were taught to follow the directions, “*push spoon*,” “*push car*,” “*point to spoon*,” “*point to car*.” They were then taught the combination “*lift feather*.” Once they responded correctly to “*lift feather*,” participants showed generalization to novel pairs (e.g., “*lift spoon*,” “*push feather*”) without direct training.

Saunders and colleagues extended this work to teaching the alphabetic principle using matrix training. The first study involved typically developing, prereading children

(Mueller, Olmi, & Saunders, 2000). The authors used a matrix composed of all combinations of three onsets and seven rimes. They taught participants to select specific printed consonant-vowel-consonant (CVC) words upon hearing a spoken-word sample. The onset is the first letter in a word (*b* in *bat*) and the rime is the vowel and remaining consonants (*-at* in *bat*).

The authors taught subsets of CVC words (e.g., *mat*, *sat*, *sop*, and *sug*), then tested accuracy on untaught words with the same units recombined (e.g., *mop* and *mug*). Selecting the correct untaught words during matching-to-sample (MTS) was considered evidence of the alphabetic principle via recombinative generalization. As with the previous studies using matrix training, once several combinations were taught, generalization to novel combinations was seen. That is, the participants were responding correctly to new words without being directly taught to do so. Later studies extended these procedures to adults with intellectual disabilities (Saunders, O'Donnell, Vaidya, & Williams, 2003; Saunders, 2011; Schmidt-Naylor, Saunders, & Brady (in press); Stewart, Hayashi, & Saunders, 2010) with positive results.

Research on Within-Syllable Units

According to the key features of instructional programming discussed earlier, complex skills are best taught in a gradual progression. When developing a program to teach abstraction of target units within spoken words, it is important to consider whether research has shown certain units to be easier to abstract. This knowledge allows us to begin with units that are easier for students to acquire, and progress to more difficult units gradually. As shown in Figure 1, spoken syllables (CVC words) can be broken into units

that are larger than phonemes, the onset and rime. The onset is the first consonant or consonant cluster of a syllable, and the rime is the vowel and remaining consonants.

Experiments using oddity tasks (Kirtley, Bryant, MacLean, & Bradley, 1989) and phoneme manipulation tasks (Treiman, 1985), have shown that students at the beginning stages of reading tend to break up spoken syllables by onset and rime units. During oddity tasks, several words are spoken to a student and s/he must identify which word does not belong (e.g. pig, hill, pin—hill does not belong because it has a different onset). A phoneme manipulation task requires a student to eliminate, re-order, or add phonemes to spoken words (Adams, 1990).

The research shows that students “find it extremely difficult to detect phonemes, except when the phoneme coincides with the word’s onset” (Goswami & Bryant, 1990, p.22). For example, it is easier for a student to break up the word *bat* into /b/ and /at/ than it is to break it into /b/ /a/ /t/. This research informed the procedures used in the current study, which began with instruction of onsets and rimes, before instruction on within-rime units (vowel and final consonant).

The current studies add to the research on the use of matrix training to teach the alphabetic principle to students in a school setting by examining the effectiveness of two consecutive steps in an instructional program. Study 1 addressed the first step, teaching recombinative generalization of onsets and rimes within CVC words. Study 2 examined the second step, teaching recombinative generalization of vowels and codas (the final consonant) of CVC words. In both studies we tested for generalization from MTS to reading trained words.

Study 1: Recombination of Onset and Rime Units

Study 1 addressed two main questions. First, will matrix training with matching-to-sample (MTS) procedures lead to acquisition of the alphabetic principle, as shown by recombinative generalization of onsets and rimes? Second, after mastery of MTS with a subset of words, will participants show generalization to reading that subset?

To answer these questions, we constructed two matrices, each organized into six, four-word subsets. We began with one matrix, systematically training and testing the six subsets one at a time (Figure 2). Following MTS training of the first subset, we tested if the participant could then read those words (Reading Generalization Test). If they failed the test, we taught reading of all four words in the subset. After a participant learned both MTS and reading the words, we tested accuracy on MTS with the next subset to be taught (MTS Generalization Test). If a participant failed to show generalization to MTS with the new subset, it was trained. This process continued until all six subsets in a matrix were mastered. Because the first subset of a matrix is always trained on MTS, there were a total of five MTS Generalization Tests and six Reading Generalization Tests for each matrix.

The dependent variables were accuracy on the MTS and the Reading Generalization Tests administered throughout Matrix Training. Generalization was said to occur if participants scored at or above 83% correct on an MTS Generalization Test, and 88% correct or higher on a Reading Generalization Test. Showing generalization to MTS with untrained subsets is considered evidence of the alphabetic principle.

We used a multiple-probe design (Horner & Baer, 1978) to show the functional relationship between our intervention and changes in accuracy on MTS and reading for all words in both matrices throughout the study. The dependent variable was accuracy on

the MTS and Reading Comprehension Tests given before and after training of each matrix. Maintaining low accuracy on the untrained matrix during the MTS and Reading Comprehension Tests following training of the first matrix would indicate that our instruction was responsible for increases in accuracy on the taught words. Once all subsets in the –ag/-ed matrix were trained, we aimed to replicate the procedures with the –it/-un matrix.

The study had six phases (Figure 2): Pretraining, baseline Comprehension Tests, training of one matrix, Comprehension Tests, training of the next matrix, and the Comprehension Tests again. Pretraining was designed to minimize errors made during Comprehension Tests due to unfamiliarity with the task, or inability to discriminate the printed words.

Method

Participants

We contacted the principals of several local elementary schools to recruit children who could identify letters, but could not read words that had not been directly taught, and whose parents would be likely to provide consent. We selected participants who could name at least 18 printed letters and read no words on the Word Attack subtest of the *Woodcock Reading Mastery Tests – Third Edition* (WRMT-III; Woodcock, 2011). Three typically developing boys were recruited. After parents signed consent forms, daily assent was obtained by asking each child, “Would you like to come work with me?”

At the onset of the study, MH was 4 years 4 months, KR was 8 years, and TS was 6 years 5 months old. At baseline, MH was in kindergarten, TS was beginning first grade, and KR was beginning second grade. Standard scores on the Peabody Picture Vocabulary

Test – 4th edition (PPVT-4) were 101 for MH, 94 for TS, and 77 for KR. All three participants read one word (the) correctly on the Word Identification subtest, and no words on the Word-Attack subtest of the WRMT-III.

Both KR and TS were receiving speech therapy during the study. Experimenters could not accurately score TS' spoken responses during the expressive letter naming assessment due to his articulation issues at the time. Thus, letter names were tested receptively; TS listened to a spoken letter name and selected from a choice pool of letters.

TS previously participated in a study during which he learned four consonant-vowel-consonant words via the same MTS task used in the current study. He was dropped four months prior to the onset of the current study due to unavailability and non-compliance during sessions. Although his MTS accuracy was relatively high at the beginning of this study, his reading accuracy was low. Thus, we included him in this study to see if he would show generalization from MTS to reading words.

Apparatus, Setting, and Stimuli

TS' sessions were completed on a Dell Inspiron Duo laptop with a 23 by 12 cm touch-sensitive monitor. All sessions with MH and KR were completed on an Acer Aspire Switch 10 laptop with a detachable 22 by 13 cm touch-sensitive monitor. Both computers had monitors that could be turned around to resemble a tablet. During MTS sessions the monitor was turned so participants could neither see, nor use, the keyboard. Participants were seated in front of the computer while the experimenter sat to the side and slightly behind to avoid providing physical cues.

For MTS sessions, Paradigm (Version 2.5.0.68) controlled session events and recorded the latencies, locations, and accuracy of all responses. Spoken words (recorded

female voice, Midwestern accent) were presented via Kidz Gear headphones. Visual stimuli were 2.5cm lowercase printed words in black Arial font. Depending on the specific trial, they were presented in the center and/or any of the four corners of the screen.

Reading sessions were conducted with PowerPoint software from the standard Microsoft Office package. Each 3cm word, in black lowercase Calibri font, was presented alone in the center of the screen on each trial. A screen-capture software program, Snagit (Version 12.4.1), was used to record the computer screen and any spoken responses made by the participant and experimenter. The experimenter controlled the presentation of each word via the arrow keys on the keyboard.

Sessions lasted no longer than fifteen minutes, with time spent on the computer ranging from approximately 3 to 8 minutes, four to five days a week. The participants in the current study could complete about 20 MTS trials in four minutes. Sessions were conducted at each child's school, at a time and location that was convenient for the child's schedule. TS and KR's sessions were conducted in the teacher's lounge, and MH's sessions were conducted in the hallway outside of his classroom. These locations were relatively quiet at the time the sessions were conducted, afternoon for TS and KR and 11am for MH.

Stimuli were consonant-vowel-consonant (CVC) words organized into two separate matrices. Each matrix contained six 4-word sets (see table 2). Each subset was designed to create all possible combinations of two rimes and two onsets (e.g. red, led, rag, lag). Both real and non-words were used. Non-words helped decrease the likelihood of participants reading some words by sight.

Procedures

The procedures described here applied to the Comprehensive Tests and Generalization Tests for both MTS and reading.

Spoken word – printed word MTS task. Each trial began with a black rectangle in the center of a white screen (Figure 3), with the spoken word repeating every two seconds. Touching the black rectangle caused it to disappear and the choice stimuli (printed words) to appear in the corners of the screen. The spoken word continued to repeat until the participant touched a printed word.

The spoken words were each presented equally often in a session, quasi-randomly. Each word served as the correct choice (S+) and the incorrect choice (S-) on an equal number of trials, and the same spoken word was presented no more than three trials in a row. Each corner on the screen served as the location for the printed word S+ and S- an equal number of times, and the same printed word did not appear in the same corner on more than three consecutive trials.

Selecting the printed word that corresponded to the spoken word produced a smiley face against a blue background for approximately 3 seconds, “Very good!” presented by the computer once, and the experimenter delivered a small edible, token, or water balloon. Incorrect responses produced a three-second time out during which the screen was black and no audio played. Both responses were followed by an inter-trial interval (ITI) of two seconds with a blank screen and no sound. A touch to the screen during the ITI had no consequence.

Instructions were provided at the start of the first trial, on the first day of MTS training. The first trial of MTS training was a prompted trial –the correct printed-word

choice was presented in the center of the screen (Figure 3). The experimenter asked, “What is the computer saying?” and the participant was required to repeat the spoken word. The experimenter instructed the child to touch the word in the center of the screen, causing the printed words to appear at the corners of the screen while the spoken word played. Without indicating which word was correct, the experimenter then said, “Look at these words. Which one is the computer saying? Touch it.” The participant was allowed to perform the task independently on the next trial and, if he executed all steps, the instructions were not provided again.

Reading task. During all reading tasks, individual words were presented successively on the computer screen. On the first trial the experimenter said, “You will see a word on the screen. Some of the words are not real words. Read each word as best as you can.” These instructions were not given after the first trial.

If the participant did not make a response within the first 5s of a trial, the experimenter said, “Make your best guess” and waited another 5s before moving on to the next trial. Correct answers produced praise and delivery of a small edible, token, or water balloon. No feedback was given for incorrect responses. The experimenter wrote down all responses on a premade scoresheet prepared prior to the session, and all sessions were recorded using Snagit (Version 12.4.1). See Appendix E for complete scoring procedures.

Interobserver Agreement. For each participant, all Reading Comprehension and Reading Generalization Tests were video recorded using Snagit (Version 12.4.1), described earlier, so they could be scored at another time by a different experimenter. No

other experimenters were present while these sessions were being scored a second time. See Appendix E for details on how experimenters were taught to score responses.

Interobserver agreement (IOA) was calculated using point-by-point agreement for at least half of the Reading Comprehension Tests and at least 33% of the Reading Generalization Tests. For each session, the number of trials on which two observers agreed was divided by the total number of trials. Average IOA scores for the Reading Comprehension Tests and the Reading Generalization Tests were calculated for each participant.

Pretraining. Pretraining was conducted to ensure participants were familiar with the MTS task using a touch-screen, could discriminate spoken samples, and could discriminate the printed-word choices.

Spoken word – picture matching. Each participant completed a two-choice MTS task during which they had to select the picture that corresponded to a spoken word. The task is the same as the Spoken word – printed word MTS task, with pictures instead of printed words.

Criterion to move on to printed-word identity matching was 4 consecutive correct responses. MH met criterion within 7 trials, and KR met criterion immediately. TS completed a similar task in a previous study, so he did not complete spoken word -picture matching in this study.

Printed- word identity matching. Participants were given an eight-trial, 4-choice identity-matching task. Each trial began with a printed word in the center of the screen. A touch to the word caused the printed words, rag, lag, red, and led, to appear at the corners of the screen, one of which matched the printed word still present in the center. Each

word served as the correct response twice. Words were presented in quasi-random order according to the same rules as the Spoken word – printed word MTS task.

On the first trial the Experimenter said, “Touch the word in the middle of the screen.” Once the participant touched the word, the experimenter said, “Good! Now touch the word that is the same as the one in the middle.” Feedback was the same as the Spoken-to-printed word MTS task.

Criterion to move on to the MTS Comprehensive Test was 88% correct (7/8 trials). MH and KR scored 88% in the first administration of the task. TS had 75% accuracy on the first administration but reached 88% correct on the second administration.

Question 1: Will participants show evidence of recombinative generalization within the MTS task with words from the trained matrix? We addressed the first question by testing accuracy on MTS with a new subset before it was trained. The dependent measure was accuracy on the MTS Generalization Test.

MTS Generalization Test. Following the MTS and reading training of a subset, all four words from the next subset to be trained were tested. All four words in a subset were presented as choices; each one was the correct choice three times, for a total of 12 trials. Feedback was as described in the Spoken word – printed word MTS task. Generalization criterion was defined as 83% correct (10/12) or higher.

Scoring less than 11/12 correct required participants to complete MTS Training. Appendix A contains all details of the training sequence. In brief, MTS training progressed through several steps, each of which required participants to discriminate words based on certain units. We began by teaching two words from a subset in a two-

choice task. We then systematically introduced the remaining words, moving from a two-choice to a three-choice task before finally training all four words at once. Participants were trained to a mastery criterion of 92% correct on the four-choice task for two consecutive days before moving on to the Reading Generalization Test for that subset.

Question 2: Will participants show evidence of generalization to reading words trained with MTS? Following the completion of MTS training with a subset, we tested for generalization to reading of the same subset via the Reading Generalization Test. Accuracy on the generalization test was the dependent variable.

Reading Generalization Test. All the words in the subset were shown twice, presented in quasi-random order. Generalization criterion was defined as 88% correct (7/8). Correct responses were followed by delivery of token, small edible, or water balloon. No feedback was given for incorrect responses and the experimenter moved on to the next trial.

If the participant did not score 100% correct, we conducted Reading Training (see Appendix B for training details). During training we presented a word on the screen, then gave the participant two spoken choices: “Does this say rag or lag?” We began by presenting choices that differed only in the onset (e.g. rag or lag), then to choices that differed only by rime (e.g. red or rag). Participants had to respond correctly to eight consecutive trials to move on to reading trials (without a choice presented by the experimenter). Participants could move on to the MTS Generalization Test for the next subset when they could read all four words correctly twice.

Measuring accuracy on each matrix. We used a multiple-probe design across matrices of CVC words to show the effect of our procedures on accuracy on both

matrices throughout the study. Our dependent measure was accuracy on the MTS and Reading Comprehension Tests, which we planned to give at three points in the study: before and after training of the -ag/-ed matrix, and after training of the -it/-un matrix.

MTS Comprehensive Test. This test included all words from both word matrices, and letters o, c, and w. Each word served as the correct choice once. Words were presented in quasi-random order and counterbalanced. Each word was presented with the other words from the same subset. The test was divided into four, 18-trial sessions that each consisted of six spoken-to-printed letter MTS trials in a three-choice format, and 12 spoken-to-printed word MTS trials in a four-choice format. Each session had an equal number of words from each matrix.

To ensure that participants understood the task, and received some access to reinforcement, two letter trials (o, c, or w) were presented at the beginning of each session, and one letter trial was presented every four trials thereafter. Two sessions, a session being one quarter of the test, could be given in a day. Feedback was the same as other MTS tasks to ensure that low accuracy was not due to the absence of feedback.

Reading Comprehension Test. The Reading Comprehension Test included all the words in the MTS Comprehensive Test, shown successively. The sequence of trials was the same as the MTS Comprehensive Test, with both letter trials and word trials included. The whole test was given in one day. If the participants gave an incorrect response, the experimenter moved to the next trial without delivering feedback. Correct responses produced verbal praise and delivery of a reinforcer, such as a small edible, water balloon, or token.

Results

The results will be discussed in four sections: performance on the baseline Comprehensive Tests, generalization to untaught words with MTS procedures (Question 1), generalization to reading a subset previously trained with MTS (Question 2), and results of all Comprehensive Tests given throughout the study.

Baseline performance on MTS and Reading Comprehensive Tests

Figure 4 shows whole-word accuracy on the MTS Comprehensive Test given before training of the –ag/-ed matrix (left bar of upper graphs). Chance level for whole word accuracy is 25% correct. TS' data are not shown because he scored 84% correct on the MTS Comprehensive Test. MH and KR's accuracy on whole words was around chance levels for each matrix. MH and KR scored above chance on onsets, however, choosing one of the two words with the correct onset on over 70% of trials. For example, if the spoken word was “pag” and the printed choices were pag, ped, tag, and ted, they selected either pag or ped.

On the baseline Reading Comprehensive Test (Figure 5), MH and TS each read one word correctly, and no other words were read correctly across the three participants. Because it is a reading task, there is no chance level of responding. As with the MTS Comprehensive Test, for MH and KR, accuracy was higher on onsets than rimes. TS showed similar, and low, accuracy across onsets and rimes. Accuracy on onsets across both matrices ranged from 13% to 50% correct, while rime accuracy ranged from 0% to 13% correct.

Question 1: Will matrix training with MTS procedures lead to acquisition of the alphabetic principle, as shown by recombinative generalization of onset and rimes?

After a participant mastered both MTS and reading of a subset, we tested for generalization to MTS with the next subset. Generalization was defined as 83% (10/12) correct on the MTS Generalization Test. Figure 6 shows accuracy on the baseline MTS Comprehensive Test (black columns) and the MTS Generalization Tests (crossed columns) for each of the six subsets in the -ag/-ed matrix. Subsets are indicated by their onsets and are presented in order of training from first (left) to last (right) along the x-axis of each graph. The first subset was taught, thus there was no generalization test.

Data for the RL subtest for TS and KR reflects rime accuracy rather than whole-word accuracy. Their onset accuracy during MTS was lowest for that subset, and their speech impairments, noted below, suggested difficulty discriminating the sounds /r/ and /l/.

Participants MH and KR scored higher on the MTS Generalization Test for each subset than on the baseline MTS Comprehensive Test. KR scored 92% correct on every MTS Generalization Test, so he received no further MTS Training after his first subset. TS had relatively high accuracy at baseline, thus he did not have as much opportunity to show generalization to MTS with untrained words.

Question 2: Will participants show generalization from MTS with a trained subset, to reading the same subset?

Figure 7 shows accuracy on the -ag/-ed matrix during the baseline Reading Comprehensive Test (black columns) and the Reading Generalization Tests (crossed columns) in the same format as Figure 6. Generalization was defined as 88% (7/8) correct and Reading Training was conducted if participants scored less than 100% correct on a generalization test.

As with the MTS task, the RL subtest for TS and KR reflects rime accuracy only, because we had difficulty scoring their spoken responses during the reading tasks for this subset. TS pronounced both /r/ and /l/ as /w/, and KR said /r/ for /l/ or did not provide an onset at all (e.g. saying “ug” for rag).

Accuracy on reading at baseline was nearly 0% correct on the –ag/-ed matrix for all participants, and all showed very large increases in accuracy on each subset during Matrix Training. MH and TS scored at least 88% correct on five and four of the six Reading Generalization Tests, respectively. KR’s accuracy ranged from 38% to 75% across subsets.

Accuracy on MTS and Reading Comprehensive Tests throughout Study 1

MTS Comprehensive Test. Figure 4 shows whole-word accuracy on the MTS Comprehensive Test given before and after training. The –ag/-ed matrix is shown in the top panel, and the –it/-un matrix is shown in the bottom panel. Data for TS and MH are the mean of two administrations of the baseline MTS Comprehensive Test. KR only had one presentation because he was refusing to come. We decided to begin KR’s training in order to increase his likelihood of remaining in the study. TS’s data are not shown because his accuracy at baseline was above 80% correct, and after matrix training he scored 92% correct on both matrices.

MH and KR scored below 40% correct on both the –ag/-ed and the –it/-un matrices during the baseline Comprehensive MTS Test. Accuracy on the –ag/-ed matrix increased following Matrix Training for both participants. However, there was also an increase in accuracy on words in the untrained, -it/-un matrix (bottom panels).

Reading Comprehensive Test. Figure 5 shows the results of the Reading Comprehensive Test given before and after training of the –ag/-ed matrix, in the same format as Figure 4. Reading accuracy at baseline was approximately 0% correct for all three participants. Following -ag/ed matrix training, all participants showed an increase in accuracy on the Reading Comprehensive Test. Accuracy increased on both the trained and untrained matrices following Matrix Training.

Although not required to do so, KR and TS produced subsyllable sounds prior to giving a whole word response during the Reading Comprehensive Test following training of the –ag/-ed matrix. We recorded the subsyllable sounds, and grouped these responses into two categories: phonemes and onset/rime responses. An example of a phoneme response is /b/ /a/ /t/, where at least two phonemes are produced and the rime is not said as a unit. An example of an onset/rime response is /b/ /at/, where the onset is said separately from the rime.

Figure 8 shows that KR was more likely to say a whole word with subsyllable sounds if the word was from the untrained (it/un) matrix. He was more likely to use onset/rime (checkered portion of columns) for words ending in –it, and phonemes (grey portion of columns) for words ending in –un. He gave subsyllable responses for all but one word from the untrained matrix.

TS almost always said the onsets and rimes for each word prior to making a whole word response. He exclusively said onset and rime for words ending in –it. Of the five words that he did not provide subsyllable sounds for, four were real words (red, bag, sun, and bed)

Trials required to reach criterion

Figure 9 shows the number of trials, both teaching and testing, required for each participant to meet criterion on MTS (black columns) and reading (crossed columns) for each subset. Subsets are shown in order of MTS training from first (left) to last (right) and are labeled by the two onsets (x-axis). If criterion was met on generalization tests, participants required only 12 trials for MTS and 8 trials for reading. MTS training added a minimum of 52 trials and reading training added at least 32 trials. Participants complete about 20 MTS trials in four minutes.

Each participant required the highest number of trials on the first two subsets, and completed his last subset in the minimum number of trials; neither MTS nor reading of that subset required training. The total number of trials required in Study 1 ranged from 488 to 596 for MTS, and from 400 to 475 for reading. Total time spent in the MTS task during the entire study ranged from approximately 1 hour 36 minutes to 2 hours.

Interobserver Agreement

IOA was calculated on half the Reading Comprehension Tests for each participant. Experimenters agreed on 99%, of trials for MH and TS, and on 100% of trials for KR. IOA was calculated on three of the Reading Generalization Tests (i.e., 50% of the tests) for MH and TS. We only calculated IOA on two of the Reading Generalization Tests (i.e., 33% of the tests) for KR, due to some videos being damaged. Mean scores for sessions with MH, TS, and KR were 100%, 96%, and 94%, respectively.

Discussion

On the baseline MTS Comprehensive Test, both MH and KR scored close to chance (25% correct), and TS scored 84% correct. MH and KR scored higher on onsets than

rimes, choosing one of the two words with the correct onset on approximately 70% of the trials.

On the baseline Reading Comprehension Test, none of the participants read more than a single word correctly. MH and KR each scored approximately 40% correct on onsets and 0% correct on rimes. TS showed less of a discrepancy between units, scoring 15% correct on onsets and 23% on rimes.

A primary question addressed in the study was whether matrix training produces recombinative generalization of onsets and rimes in CVC words. All participants met the MTS generalization criteria with untrained subsets within the –ag/-ed matrix, thus demonstrating the alphabetic principle via recombinative generalization. For instance, learning rag and led, then selecting red and lag without further training is evidence of the alphabetic principle. In this case, it shows participants can identify and recombine units within words, the onsets (r and l) and the rimes (ag and ed).

However, we cannot conclude that our training alone was responsible for the increase in accuracy on the –ag/-ed matrix on both the MTS and Reading Comprehension Tests, because all participants showed an increase in accuracy on the untrained, -it/-un matrix as well. Accuracy on subsyllable units throughout the study suggest that the increase in the untrained matrix on the Comprehension Tests was related to the words used in the study. In particular, participants could respond correctly based on the last letter in the rime (coda), and there were consonants that appeared in both matrices.

The matrices were not designed to draw attention to the units within each rime, because we wanted to encourage recombinative generalization of onsets and rimes as units. Thus, in both matrices, each coda was combined with only one vowel. For

example, vowel a and coda g were always paired (-ag), but vowel a and coda d were never paired (-ad). Because of the unique vowel-coda pairs, participants could respond correctly based on whole rimes or the consonants alone. During Reading tasks, most errors were due to producing an incorrect vowel, suggesting they were responding based on consonants.

In addition to responding based on consonants, there is also evidence of participants responding to rimes as a unit. In particular, participants may have read the rime, -it, by sight because it is a real word. TS and KR produced subsyllable sounds prior to whole word responses on nearly all trials during the Reading Comprehension Test following -ag/-ed training. Analysis of these responses showed TS and KR were more likely to give onset/rime responses when presented with words ending in -it, than with the other rimes. Following -ag/-ed matrix training, participants could abstract rime units across words, and recombine those units. If they could read the word *it* by sight, as the data suggest, it is likely they would apply these skills to the words ending in -it.

Related to this argument, having some consonants appear in both matrices may also have led to an increase in accuracy in the untrained matrix. One of the rimes in the untrained matrix, -un, contained the consonant n, which also served as an onset in the trained matrix (the NJ subset). It is likely that abstraction of the phoneme, /n/, at the onset of words during training generalized to the end of words during the MTS and Reading Comprehension Tests. Byrne found that “Preschool children can be taught to recognize the identity of phonemic segments [...] and the position of the target phonemes in the word appears to make no difference” (1990, p. 810).

The second major question addressed in the study was whether, after demonstrating high accuracy on MTS with a subset, participants would show generalization to reading the same subset. All participants showed higher accuracy on the Reading Generalization Test for each subset than on the Reading Comprehension Test.

Two of the three participants maintained high accuracy in the final Reading Comprehension Test. MH was the exception. Although he scored at or above 88% correct on five of the six Reading Generalization Tests on the final Reading Comprehension Test his mean score on the –ag/-ed matrix was 54% correct.

MH's reading accuracy on the trained words may have decreased during the Reading Comprehension Test because the test contains four rimes. Both the MTS and Reading Comprehension Test included all words from the –ag/-ed and the –it/-un matrices. That is, the tests contain two vowels that MH had limited exposure to, i and u. Each Reading Generalization Test followed MTS training/testing of a subset, making it likely that MH's responses on those tests were limited to the two rimes from the preceding MTS task, essentially turning the reading task into a two-choice task. Mixing all four vowels within the Comprehension Tests may have disrupted stimulus control by the two trained rimes, -ag and -ed.

Analyses of MH's responses during the Reading Comprehension Test support inconsistent stimulus control by vowels. Results of the Reading Comprehension Tests show that, for both the trained and untrained matrices, MH's onset and coda accuracy increased after –ag/-ed training. Although vowel accuracy for both matrices was substantially higher than the vowel accuracy at baseline (6% correct), it was only 46%. He said the vowel e for words containing the letter e, but also for the majority of words

containing a and u. The reason for MH's low accuracy on vowels is that MTS training in Study 1 was not designed to force attention to vowels.

Study 2: Recombination of Vowel and Coda Units

In Study 1, training was designed to promote recombinative generalization of onsets and rimes. Study 2 is the next step of instructional programming, designed to teach recombinative generalization of vowels and codas (the last consonant) within CVC words.

Method and Results

Study 2 asked the same questions as Study 1, using the same design. The methods were identical to those in Study 1 with the following exceptions: the matrices recombined vowels and codas making four rimes per subset, the Reading Comprehension Test was always given prior to the MTS Comprehensive Test, and Subset Reading Reviews were given throughout to promote maintenance of reading accuracy on trained words.

Participants

MH and KR from Study 1 participated in Study 2.

Stimuli

Forty-eight CVC words were organized into two matrices (table 3). Each matrix contained all possible combinations of four rimes and six onsets, making 24 words. Each matrix was further broken down into six subsets, composed of one onset and four rimes (e.g. rag, red, rad, reg). Only one onset was used for each subset because the focus was teaching participants to attend to the units within the rimes. The rimes included all four combinations of the two vowels and two codas in each matrix.

None of the rimes were real words (e.g. –it or –is). There were no mirror image letters (e.g. b and d) within the same matrix, and onsets and codas within a subset were never identical (e.g. mam). To decrease the likelihood of reading words by sight, each matrix included a maximum of five real words. All remaining stimuli were non-words. Subsets will be referred to by the onset (e.g. J subset). Matrices will be referred to by the vowels (e.g. i/u matrix and a/e matrix).

Procedures and Results

The procedures and results will be discussed in the following order: Pretraining, baseline Comprehensive Tests, the Subset Reading Reviews, the MTS and Reading Generalization Tests (Questions 1 & 2), and all the Comprehensive Tests given throughout the study.

Printed- word identity matching

The methods were identical to those used in Study 1, except the words rag, red, reg, and rad were used to ensure discrimination of individual letters within the rime (see Figure 1). Both participants scored above the 88% correct criterion to move on to the baseline Comprehensive Tests.

Baseline Reading and MTS Comprehensive Tests

Figure 11 shows the results of the Reading Comprehensive Test given before and after training of a single matrix. At baseline MH made near perfect scores on onsets and codas from each matrix, but scored 0% correct on a/e vowels, and 25% correct on i/u vowels. KR's accuracy was generally high across all positions in the words from each matrix, but he scored slightly higher on codas from the a/e matrix (100% correct) than the i/u matrix (79% correct).

Figure 12 shows the results of the MTS Comprehensive Test given before and after training of a single matrix. At baseline MH had two administrations of the MTS Comprehensive Test, and KR had one. Both scored approximately 55% correct on whole words for each matrix, and scored higher on codas than vowels. MH scored nearly 50% correct on vowels from each matrix, but 100% correct on codas. On vowels, KR scored 63% and 58% correct on the a/e and i/u matrices respectively. He scored 92% and 96% correct on codas from the a/e and i/u matrices respectively.

Subset Reading Reviews

After completing all training and testing for a subset, and before testing for generalization to MTS with a new subset, participants were given a reading review of previously mastered subsets. The reviews were given in a staggered fashion, with the first one given after mastery of *the second* subset. The first review included words from the first subset only. The second was given after the mastery of the third subset, and included words from the first two subsets trained, and so on.

Words were presented three times each in quasi-random order. The first and second reviews contained 12 and 24 trials, and the rest of the reviews contained 36 trials. Once a subset had been reviewed three times, it was dropped to allow for another subset to be reviewed without adding trials. Correct responses were followed with either a token or an edible. If a participant responded incorrectly, the experimenter told the participant the correct answer and the participant repeated it correctly once. Both participants reliably repeated the experimenter correctly.

Criterion to move on to the MTS Generalization Test for the next subset was 92% correct on a Subset Reading Review. Both participants scored above 92% correct on each of the five Subset Reading Reviews given.

Question 1: Will recombinative generalization occur within a matrix during MTS?

Because both participants scored approximately 50% correct on both matrices during the baseline MTS Comprehensive Test, we chose to begin MTS Training with the matrix on which participants showed the lowest accuracy during the Reading Comprehensive Test. This provided the most opportunity for each participant to show generalization to reading.

MH was trained on the a/e matrix, and KR was trained on the i/u matrix. Each participant began MTS Full Training with the subset, from the chosen matrix, on which he showed the highest accuracy on the MTS Comprehensive Test. Because the subsets in the new matrices each contain a single onset and four rimes, the procedures for MTS training varied slightly from those used in Study 1.

Appendix C contains a complete description of the steps in instructional programming). In brief, we progressed through a two-choice task, to a three-choice task, before requiring participants to complete the four-choice task. Each step required that participants discriminate between words based solely on a certain position (e.g. presenting two words that differ only in the vowels). Before moving on to the Reading Generalization Test for a subset, participants had to show 92% accuracy on the four-choice task for two consecutive days.

MTS Generalization Test. As in Study 1, each MTS Generalization Test was given following mastery of the previous subset. The first subset was not tested for

generalization because it was trained. Generalization was defined as 83% correct (10/12). If participants failed to show generalization, they were trained to a mastery criterion of 92% correct on the four-choice task for two consecutive days (see Appendix C for training details).

Figure 12 shows accuracy for each subset on the baseline MTS Comprehensive Test, and the MTS Generalization Tests. Subsets are shown across the x-axis in order of training, and are indicated by their onset. Both participants showed higher accuracy on the Generalization Test for each subset than on the baseline MTS Comprehensive Test. MH reached MTS generalization criterion on three of the four tests he was given. He was not given the MTS Generalization Test for the N subset due to high accuracy on an added reading probe, described below. KR reached criterion on four of the five MTS Generalization Tests.

Question 2: Will participants show generalization from MTS with a trained subset, to reading the same subset?

Reading Generalization Test. Following the completion of the MTS Generalization Test and/or MTS Training for each subset, we tested reading of that subset. Generalization from MTS to reading a subset was defined as scoring 88% correct or higher (7/8 correct) on the Reading Generalization Test. If a participant scored less than 100%, he received reading training.

Scoring 6 or 7 correct on the Reading Generalization Test required a participant to complete Error Contrast Reading Training, and scoring 5 or less correct required a participant to complete Verbal Comparison Training (see Appendix D for full training details). Error Contrast Reading Training includes only those words that the participant

made an error on. For instance, if a participant said rag when presented with red, he would receive reading training on just those two words. Similar to Study 1, participants were presented with a word and given a choice by the experimenter, “Does this say red or rag?” Verbal Comparison Training was in the same format, but we systematically worked through training all the words in a subset. Before moving on the MTS Generalization test for the next subset, participants had to read all words in a subset correctly twice in a row.

Figure 13 shows the results of the baseline Reading Comprehensive Test (black columns), and the Reading Generalization Tests (crossed columns), in the same format as Figure 12. MH went from reading none of the words from the a/e matrix at baseline, to scoring 100% correct on four of the six Reading Generalization Tests. KR showed generalization to reading on four of the five subsets for which he was given a generalization test. Results for the G subset are not shown on the graph because he did not receive MTS training for that subset due to time constraints.

Accelerated procedures for MH. We modified the procedures to be more time efficient for MH because the end of the school year was approaching and his accuracy was high on the first three Reading Generalization Tests. He was given a Reading Probe after mastery of the third subset. The probe consisted of the Reading Generalization Tests for the final three subsets given successively in the same session.

Scoring less than 100% on a subset in the probe meant that training continued as planned, with the MTS Generalization Test for that subset, followed by either MTS Training or the Reading Generalization Test. MH scored 100% correct on the N subset, and no further training was administered for that subset. He scored 7/8 correct on the M and H subsets.

Following the Reading Probe, MH took the MTS Generalization Test for the M subset. Data shown for subsets M and H in Figure 13 reflect accuracy on the Reading Generalization Tests following MTS (not the probe data).

Accuracy on Reading and MTS Comprehensive Tests

Reading Comprehensive Test. Figure 10 shows accuracy on the Reading Comprehensive Test before and after training. MH was trained on the a/e matrix, and KR was trained on the i/u matrix. Both participants showed an increase in accuracy on both matrices (trained and untrained) following Matrix Training.

As in Study 1, KR produced subsyllable sounds prior to whole-word responses during the Reading Comprehensive Test. Figure 14 shows KR's subsyllable responses on the Reading Comprehensive Test before and after training of the i/u matrix. Before training, KR produced individual phonemes more often than onset/rime units prior to giving a whole-word response, and showed the opposite after training.

MTS Comprehensive Test. For both MH and KR, accuracy on *each* matrix during the test given at baseline was approximately 55% correct. Both participants showed a large increase in accuracy following training (Figure 11), for both the trained and untrained matrices.

Trials required to reach criterion

Figure 15 shows the number of trials required for MH (left) and KR (right) to reach criterion for both MTS (black columns) and reading (crossed columns) for each subset. Subsets, indicated by the onset, are shown across the x-axis in order of training from left to right. It takes about 4 minutes to complete 20 MTS trials.

Both participants required the minimum number of trials, 12, to reach criterion for MTS on two of the six subsets taught, and over 150 trials to reach criterion on MTS for each of the remaining four subsets. Both participants required only 8 trials to reach criterion on reading for four of six subsets (i.e., they demonstrated generalization to reading). The total number of trials required in Study 2 ranged from 563 to 880 for MTS, and from 184 to 60 for reading. Total time spent in the MTS task during the entire study ranged from approximately 1 hour 50 minutes to 3 hours.

Interobserver Agreement

IOA was calculated on both of the Reading Comprehension Tests for MH and on one of the two Reading Comprehension Tests for KR. Observers agreed on 93% and 99% of trials for MH and KR, respectively. IOA was calculated on the first three Reading Generalization Tests (i.e., 50% of the tests) for both participants. Observers agreed on 100% of trials for both MH and KR.

Discussion

Accuracy on the MTS Comprehensive Test at baseline was approximately 50% on each matrix for both participants. Accuracy was high on consonants and low on vowels. This was expected because Study 1 did not require participants to attend to vowels. Results were more variable on the baseline Reading Comprehension Test. Both participants showed high accuracy on consonants, but accuracy on vowels was low for MH and relatively high for KR.

Study 2 asked whether matrix training would produce recombinative generalization of vowels and codas during the MTS task, thus demonstrating the alphabetic principle. Both participants showed generalization to untaught subsets in the trained matrices (a/e matrix

for MH, and i/u matrix for KR). That is, they recognized sound-letter relations across words, even though the words were not trained. For instance, learning rag and red, then selecting reg and rad without further training is evidence of the alphabetic principle.

Participants also showed strong evidence of generalization from MTS to reading subsets from the trained matrices, another question addressed in Study 2. MH showed generalization on five out of the six subsets, and KR showed generalization on all five subsets tested.

Compared to baseline, accuracy on the trained matrix increased on the MTS Comprehensive Test for both participants. The same was seen on the Reading Comprehensive Test. However, we cannot conclude that our training, as opposed to outside instruction, was solely responsible for the increase because both participants also showed an increase on the untrained matrix during the MTS Comprehensive Test. Additionally, MH showed an increase in accuracy on the untrained matrix during the Reading Comprehensive Test.

It is plausible that aspects of our procedures produced the increased accuracy on the untrained matrix during the MTS and Reading Comprehensive Tests. As stated earlier, both participants showed the alphabetic principle via generalization to MTS with untaught subtests within the trained matrix. Once the alphabetic principle is demonstrated with some letters, children can demonstrate it more rapidly with additional letters. Byrne & Fielding-Barnsley (1990) showed this type of rapid generalization with typically developing, prereading children. In our study, although each word appeared only once in the Comprehensive Test, and accuracy was thus assessed on each word prior to feedback, feedback during the Comprehensive Tests could have been enough to promote

generalization across all words in the untrained matrix. That is, once MH and KR demonstrated generalization in the trained matrix, they demonstrated generalization (particularly the vowels) in the untrained matrix with only the feedback given during the Comprehensive Tests.

KR showed high accuracy on vowels during the baseline Reading Comprehensive Test, which was unexpected because Study 1 did not teach participants to attend to vowels. One reason that his vowel accuracy was high may be related to the subsyllable responses he provided prior to his whole-word response. These subsyllable responses provided by KR may have acted as a differential observing response (DOR). “Differential observing response procedures control observing behavior and verify discrimination of critical stimuli or stimulus features” (Walpole, Roscoe, & Dube, 2007). In this case, KR attended to each letter in the words via his subsyllable responses. It is likely that this behavior led to higher accuracy.

To find out if KR’s reading accuracy was high because of this DOR, we probed KR’s accuracy when he was given a limited time to respond. The probe was given following the Reading Comprehensive Test after training the i/u matrix. It was identical to the Reading Comprehensive Test, except that KR had 2s after a word was presented to give a spoken response. He was instructed to say each word without trying to sound it out beforehand. On each trial, the printed word was shown in the center of the white screen, one second later a black box replaced the word, and another second later the box disappeared and the screen turned blue. Responses given in the presence of the blue screen were recorded by the experimenter, but were considered incorrect. This procedure blocked KR from engaging in the DOR, and his accuracy went down to 0% correct.

General Discussion

The current studies add to previous research by using a computerized instructional program to teach the alphabetic principle to typically developing children who were recommended by teachers as being behind their peers. The studies were designed to act as two steps in an instructional program, each designed to promote recombinative generalization of subsyllable units via matrix training. Study 1 focused on onsets and rimes, and Study 2 on vowels and codas. In both studies, all participants acquired the alphabetic principle, as shown via accuracy on the MTS Generalization Tests given during Matrix Training. They also showed generalization from spoken-to-printed word MTS to reading.

We cannot definitively conclude that our training was solely responsible for the increase in accuracy on the Comprehensive Tests due to an increase in the untrained (control) matrix in each study. However, we propose that the results the MTS Comprehensive Tests across studies can be interpreted as a multiple baseline across the two phonemic units, consonants and vowels.

The Comprehensive Tests given at the onset of Study 2 be interpreted as an additional untrained baseline for Study 1, in that participants had not yet been taught to attend to the vowels within words. During the Study 2 baseline MTS Comprehensive Test, both participants selected one of the two choices with the correct coda, demonstrating stimulus control by the consonants, which is what we taught in Study 1. Vowel accuracy was around chance levels. Matrix training in Study 2 was designed to produce recombination of vowels and codas, and led to stimulus control by vowels and consonants, which was shown in the Study 2 MTS Comprehensive Test after training.

MH's Reading Comprehension Tests show further evidence that our procedures taught stimulus control by consonants in Study 1. MH received additional reading training at the end of Study 1, before receiving the Reading Comprehension Test for Study 1 again, scoring 100% correct. Despite his perfect accuracy following the extra training, MH's responses on the Study 2 baseline Reading Comprehension Test were only 25% correct on vowels. These responses support the notion that using words with unique vowel-coda pairs for the rimes led to control by consonants, or the rime as a whole, as opposed to consonants *and* vowels.

Future studies should test reading accuracy on the individual rimes at the beginning of each study. Doing so here may have shown higher accuracy on certain rimes. In particular, we may have found that participants could read the rime –it at the onset of Study 1, and thus we would have changed the word matrices so that rime was not included.

It is also encouraged that future studies not include reading tasks until after MTS instruction at the phoneme level. In Study 1, participants required more reading training than MTS training because the MTS task did not prepare them to attend to every phoneme in the words. The low vowel accuracy on the Reading Comprehension Test given at the onset of Study 2 indicates that there was not stimulus control at the phoneme level. That is, participants were not at the point in instruction where they should be expected to produce all sounds in the words at high accuracy. Instead, we saw that they could accurately produce consonants, but not vowels. In Study 2, participants required more MTS training than reading training. The fact that very little reading training was required suggests that the MTS tasks had adequately prepared them for the reading tasks

by establishing stimulus control by all letters in a word, therefore leading to high accuracy on the Reading Generalization Tests and less reading training.

The overall goal of our research is to create an instructional program to teach pre-reading skills to children with a range of prerequisite skills. Thus, future studies should consider additional steps that will be necessary in teaching pre-reading children. For instance, there is the possibility that the participants' high accuracy on onsets made the MTS task easier by turning a four-choice task into a two-choice task. That is, the participants could exclude the two choices that had the incorrect onset, therefore leaving only two words to choose from. It is unclear if the procedures will be effective with those who don't demonstrate some knowledge of onsets. The participants could also discriminate between the printed words, as shown during pretraining. Working with children with fewer prerequisite skills will provide information on the additional steps to be added to the overall program.

The current studies offer further support of the use of technology for instructional programming and research in general. Using a computer program allowed us to present many trials in a short time (about 5 trials per minute), and the touch-screen monitor provides immediate feedback for each response. There is also the inherent procedural integrity of computerized tasks because there is very little room for experimenter error. The compact computers and the headphones allowed us to work with children in environments that were not conducive to tabletop instruction.

References

- Adams, M. J. (1990). *Beginning to read: Learning and thinking about print*. Cambridge, MA: MIT Press.
- Axe, J. B., & Sainato, D. M. (2010). Matrix training of preliteracy skills with preschoolers with autism. *Journal of applied behavior analysis*, 43(4), 635-652.
- Byrne, B., & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology*, 81(3), 313-321.
- Byrne, B., & Fielding-Barnsley, R. (1990). Acquiring the alphabetic principle: A case for teaching recognition of phoneme identity. *Journal of Educational Psychology*, 82(4), 805.
- Dunn, L. M., Dunn, L. M., Bulheller, S., & Häcker, H. (1965). *Peabody picture vocabulary test*. Circle Pines, MN: American Guidance Service.
- Catania, A. C. (1998). *Learning 4th edition*. Englewoods Cliffs, NJ: Sloan Publishing.
- Goldstein, H. (1983). Recombinative generalization: Relationships between environmental conditions and the linguistic repertoires of language learners. *Analysis and Intervention in Developmental Disabilities*, 3(4), 279-293.
- Goldstein, H., & Mousetis, L. (1989). Generalized language learning by children with severe mental retardation: effects of peers' expressive modeling. *Journal of applied behavior analysis*, 22(3), 245-259.
- Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hove, East Sussex: Psychology Press Ltd.

- Holland, J. G. (1960). Teaching machines: an application of principles from the laboratory. *Journal of the Experimental Analysis of Behavior*, 3(4), 275-287.
- Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: a variation of the multiple baseline. *Journal of applied behavior analysis*, 11(1), 189-196.
- Keller, F. S. (1968). "good-bye, teacher...". *Journal of applied behavior analysis*, 1(1), 79-89.
- Kirtley, C., Bryant, P., MacLean, M., & Bradley, L. (1989). Rhyme, rime, and the onset of reading. *Journal of experimental child psychology*, 48(2), 224-245.
- Mineo, B. A., & Goldstein, H. (1990). Generalized learning of receptive and expressive action-object responses by language-delayed preschoolers. *Journal of Speech and Hearing Disorders*, 55(4), 665-678.
- Mueller, M. M., Olmi, D. J., & Saunders, K. J. (2000). Recombinative generalization of within-syllable units in prereading children. *Journal of Applied Behavior Analysis*, 33(4), 515-531.
- National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.
- Saunders, K. J., O'Donnell, J., Vaidya, M., & Williams, D. C. (2003). Recombinative generalization of within-syllable units in nonreading adults with mental retardation. *Journal of Applied Behavior Analysis*, 36(1), 95-99.
- Skinner, B.F. (1957). *Verbal behavior*. Prentice-Hall, Inc.

- Saunders, K. J. (2011). Designing instructional programming for early reading skills. *Handbook of applied behavior analysis*, 92-109.
- Schmidt-Naylor, A. C., Saunders, K.J., & Brady, N.C. (in press). Developing the alphabetic principle to aid text-based AAC use by adults with low speech intelligibility and intellectual disabilities. *American Journal of Speech Language Pathology*.
- Schuele, C. M., & Boudreau, D. (2008). Phonological awareness intervention: Beyond the basics. *Language, Speech, and Hearing Services in Schools*, 39(1), 3-20. Retrieved from <http://search.proquest.com/docview/622082622?accountid=14556>
- Snagit (Version 12.4.1) [Computer software]. Okemos, MI: TechSmith Corp. (2015)
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. National Academies Press.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading research quarterly*, 360-407.
- Stewart, K., Hayashi, Y., & Saunders, K. (2010). Enhancing vowel discrimination using constructed spelling. *The Analysis of verbal behavior*, 26(1), 57.
- Striefel, S., & Wetherby, B. (1973). Instruction-following behavior of a retarded child and its controlling stimuli. *Journal of Applied Behavior Analysis*, 6(4), 663-670.
- Striefel, S., Wetherby, B., & Karlan, G. R. (1976). Establishing generalized verb-noun instruction-following skills in retarded children. *Journal of Experimental Child Psychology*, 22(2), 247-260.
- Tagliaferri, B. (2015). Paradigm (Version 2.5.0.68) [Computer software]. Perception Research Systems Inc.

- Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children. *Journal of experimental child psychology*, 39(1), 161-181.
- U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2015 Reading Assessment.
- Walpole, C. W., Roscoe, E. M., & Dube, W. V. (2007). Use of a differential observing response to expand restricted stimulus control. *Journal of Applied Behavior Analysis*, 40(4), 707-712.
- Wetherby, B., & Striefel, S. (1978). Application of miniature linguistic system or matrix training procedures. *Language intervention strategies*, 319-356.
- Whitehurst, G. J. (1971). Generalized labeling on the basis of structural response classes by two young children. *Journal of Experimental Child Psychology*, 12(1), 59-71.
- Woodcock, R.M. (2011). *Woodcock Reading Mastery Tests – 3rd ed.* Bloomington, MN: NCS Pearson, Inc.

List of Illustrative Materials

Figure Captions

- Figure 1.* The parts of a syllable, *bat*. 52
- Figure 2.* The six phases of Study 1, shown in order from top to bottom. 53
- Figure 3.* The trial layout for the spoken-to-printed word MTS task (top), 54
prompted MTS trials (see Appendix), and the printed- word identity
matching (bottom) are shown as progressing from left to right.
- Figure 4.* The percentage of correct responses made by MH (left) and KR (right) 55
during the MTS Comprehensive Test before and after training of the -
ag/-ed matrix. The rimes at the top left of each panel indicate the matrix
represented. The black line represents –ag/-ed training.
- Figure 5.* The percentage of correct responses made by MH (left), TS (middle), 56
and KR (right) during the Reading Comprehensive Test are shown. The
vertical black line represents -ag/-ed Matrix Training.
- Figure 6.* The percentage of correct responses made by MH (top), TS (middle), 57
and KR (bottom) on the baseline MTS Comprehensive Test (black
columns) and the MTS Generalization Tests (crossed columns) given
during Matrix Training for each subset in the -ag/-ed matrix. Subsets
are shown in order of training from left to right on the x-axis.
- *Accuracy on the RL subset for TS and KR are based on rimes.
- Figure 7.* The percentage of correct responses made by MH (top), TS (middle), 58
and KR (bottom) on the baseline Reading Comprehensive Test (black
columns) and the Reading Generalization Tests (crossed columns)

given during Matrix Training for each subset in the -ag/-ed matrix.

Subsets are shown in order of training from left to right on the x-axis.

*Accuracy on the RL subset for TS and KR are based on rimes.

Figure 8. The number of whole word responses (out of 12 opportunities) that were preceded by subsyllable sounds (y-axis) during the Reading Comprehensive Test given following mastery of the -ag/-ed matrix. TS is shown on the top panel, and KR on the bottom. The rime units of each matrix are presented across the x-axis. Each column shows the number of whole word responses that were preceded by phonemes (grey) and onset/rime sounds (checkered). This figure does not reflect accuracy of responses. 59

Figure 9. The number of trials (y-axis) needed for MH (top), TS (middle), and KR (bottom) to reach mastery criteria on MTS (black columns) and Reading (crossed columns) for each subset (x-axis). Subsets are indicated by their onsets and are shown in order of training from left to right. *Accuracy on the generalization tests for the RL subset for TS and KR were based on rimes only. 60

Figure 10. Accuracy on the Reading Comprehensive Test before and after training for MH (left) and KR (right). Matrices are indicated by the vowels at the top left of each panel. Training is indicated by a black vertical line. 61

Figure 11. Accuracy on the MTS Comprehensive Test before and after training for MH (left) and KR (right). Matrices are indicated by the vowels at 62

the top left of each panel. Training is indicated by a black vertical line.

Figure 12. Accuracy on the baseline MTS Comprehensive Test (black columns) 63

and the MTS Generalization Test (crossed columns) for each subset, indicated by the onset across the x-axis, for MH (top) and KR (bottom). The arrow in MH's graph represents when the Reading Probe was given. *There was not an MTS Generalization Test for the N Subset.

Figure 13. The percentage of correct responses made by MH (top) and KR 64

(bottom) on the baseline Reading Comprehensive Test (black columns) and the Reading Generalization Tests (crossed columns) given during Matrix Training for each subset trained in Study 2. Subsets are shown in order of training from left to right on the x-axis. The arrow in MH's graph represents when the Reading Probe was given (see Appendix).

*Due to time constraints, KR did not receive MTS training on the G subset, therefore his Reading Generalization Test score for that subset is not shown.

Figure 14. The number of KR's whole word responses (out of six opportunities) 65

that were preceded by subsyllable sounds during the Reading Comprehensive Test given before (top) and after (bottom) training. The rime units of each matrix are presented across the x-axis. For each rime there is a column showing the number of responses to words ending in that rime that were preceded by phonemes (grey), onset/rime sounds (checkered). This figure does not reflect accuracy of responses.

Figure 15. The number of trials (y-axis) needed for MH (top) and KR (bottom) to 66

reach mastery criteria on MTS (black columns) and Reading (crossed columns) for each subset (x-axis). Subsets are indicated by their onsets and are shown in order of training from left to right. *G subset was not trained for reading due to time constraints.

Figure 16. A scoresheet, with added column for “Nature of Error,” used to show the possible spoken responses a participant can give and how the experimenter should record them during the session.

Table 1

Participant Demographics

	Study	Age Yr; mo	Speech Therapy	Numbers of Letters Identified	Peabody Picture Vocabulary Test Standard Score	Woodcock Reading Mastery Test III Raw Scores	
						Word ID	Word Attack
MH	1 & 2	4;4	N	26	101	1	0
TS	1	6;5	Y	20*	94	1	0
KR	1 & 2	8	Y	23	77	1	0

*TS had to identify letters receptively instead of having to name the letters.

Table 2

The word matrices used in Study 1.

-ag/-ed Matrix				-it/-un Matrix			
rag	lag	red	led	fit	kit	fun	kun
pag	tag	ped	ted	mit	sit	mun	sun
nag	jag	ned	jed	hit	bit	hun	bun
wag	sag	wed	sed	rit	jit	run	jun
mag	hag	med	hed	pit	lit	pun	lun
kag	bag	ked	bed	git	dit	gun	dun

Table 3

The word matrices used in Study 2.

a/e Matrix				i/u Matrix			
jag	jeg	jed	jad	gim	gum	gic	guc
kag	keg	ked	kad	rim	rum	ric	ruc
hag	heg	hed	had	lim	lum	lic	luc
pag	peg	ped	pad	tim	tum	tic	tuc
nag	neg	ned	nad	bim	bum	bic	buc
mag	meg	med	mad	wim	wum	wic	wuc

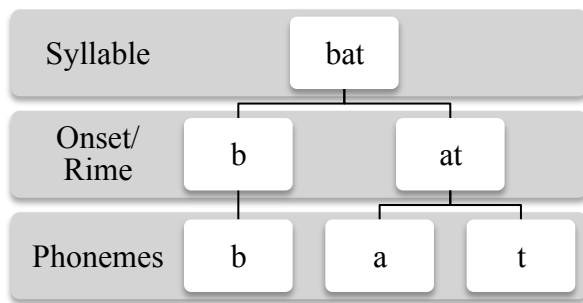


Figure 1. The parts of a syllable, *bat*.

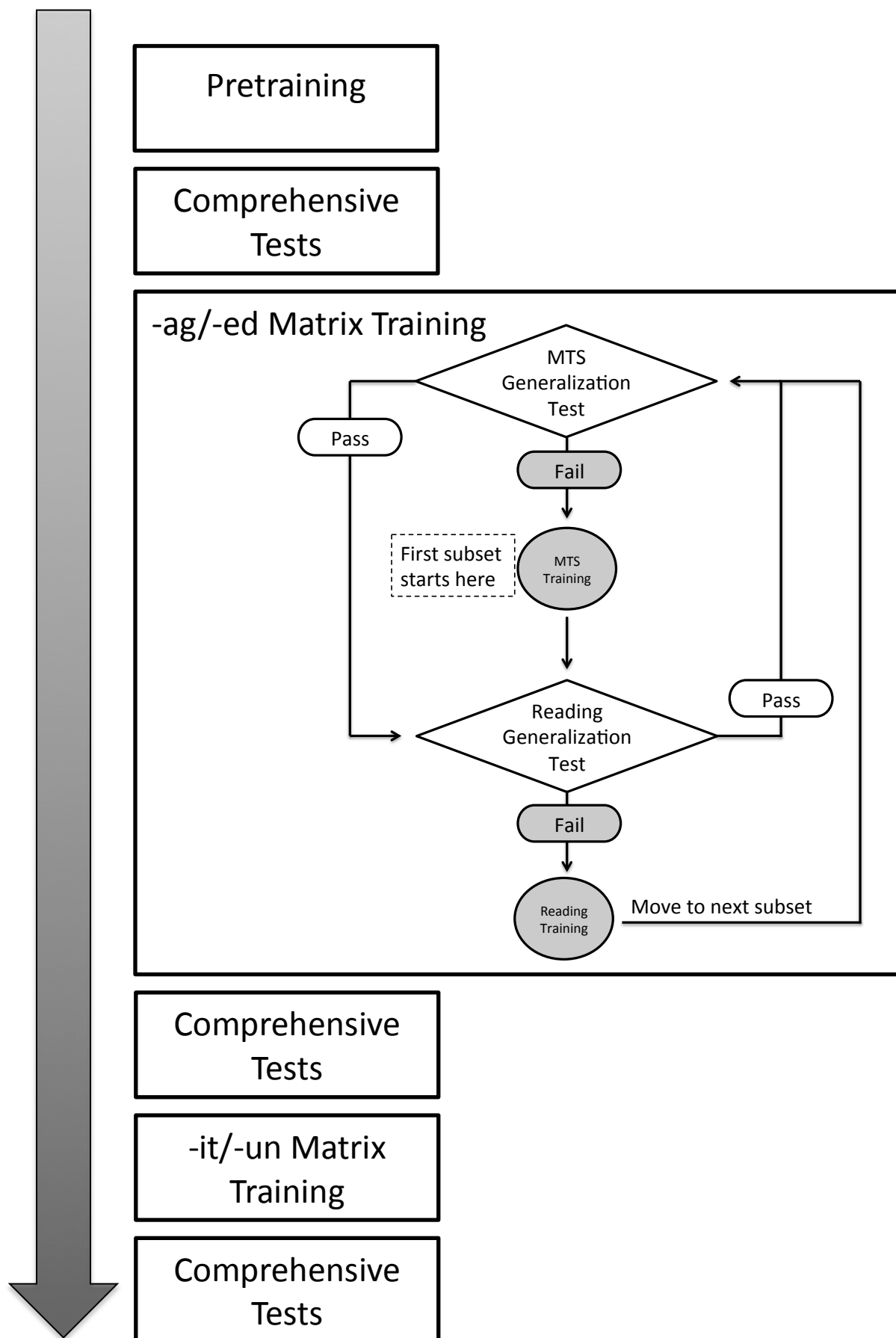
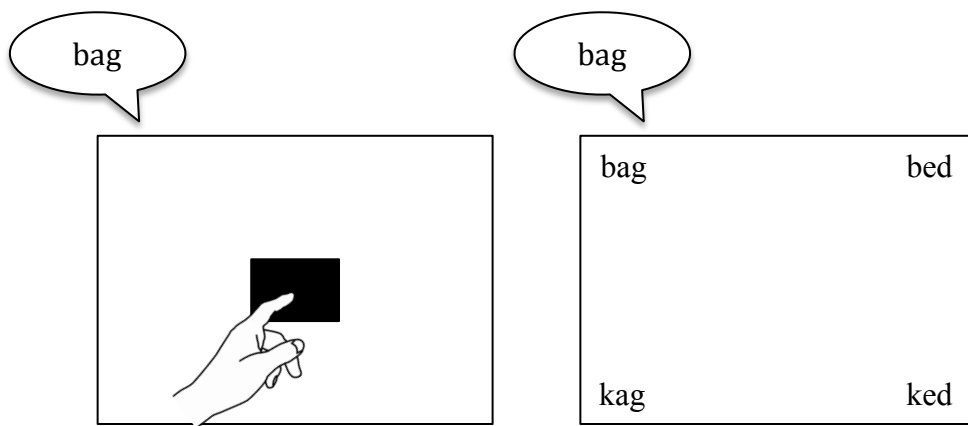
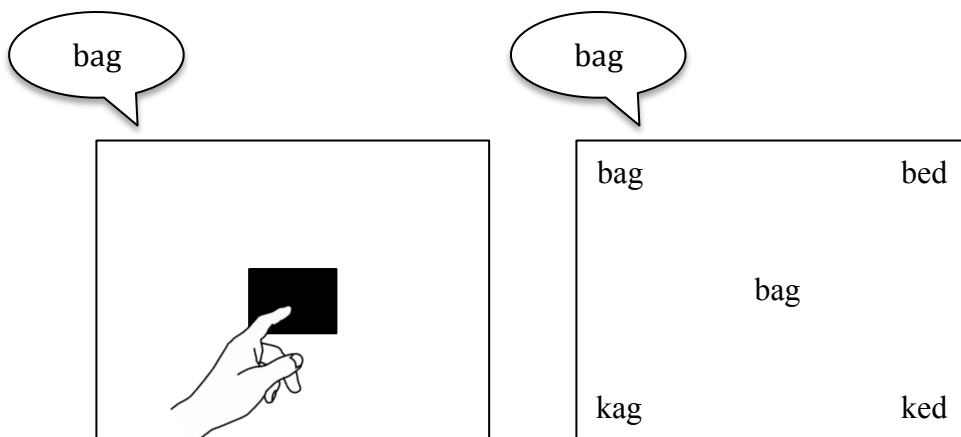


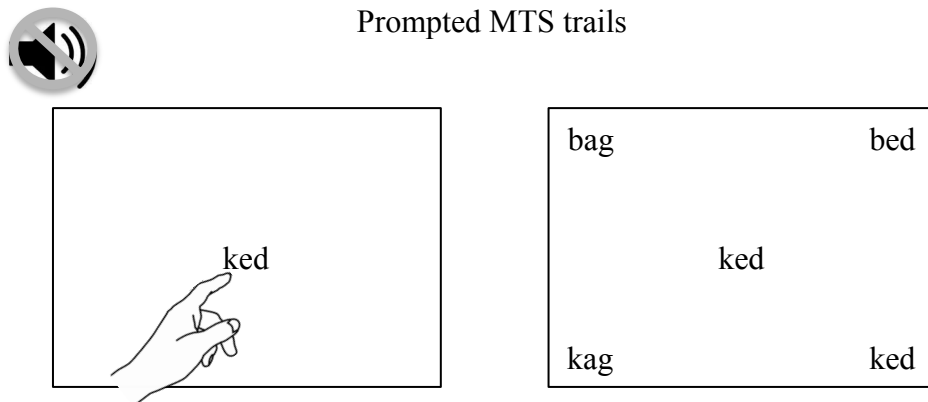
Figure 2: The six phases of Study 1, shown in order from top to bottom.



Spoken-to-printed word MTS task

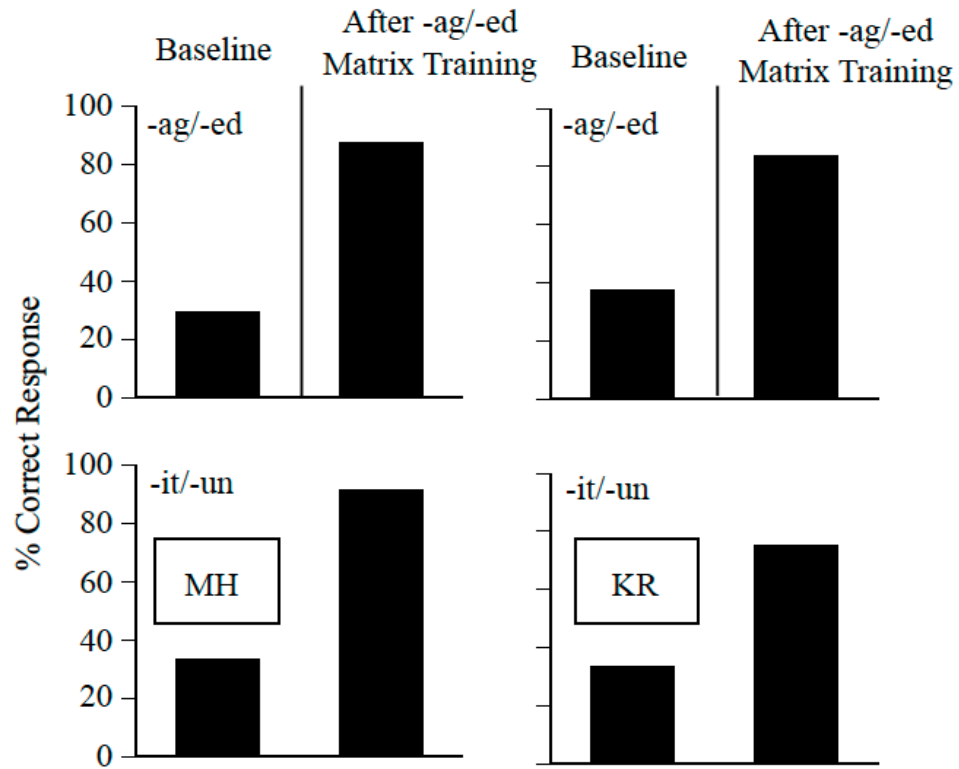


Prompted MTS trails



Printed-word Identity Matching

Figure 3. The trial layout for the spoken-to-printed word MTS task (top), prompted MTS trials (see Appendix A), and the printed- word identity matching (bottom) are shown as progressing from left to right.



MTS Comprehensive Test

Figure 4. The percentage of correct responses made by MH (left) and KR (right) during the MTS Comprehensive Test before and after training of the -ag/-ed matrix. The rimes at the top left of each panel indicate the matrix represented. The black line represents – ag/-ed training.

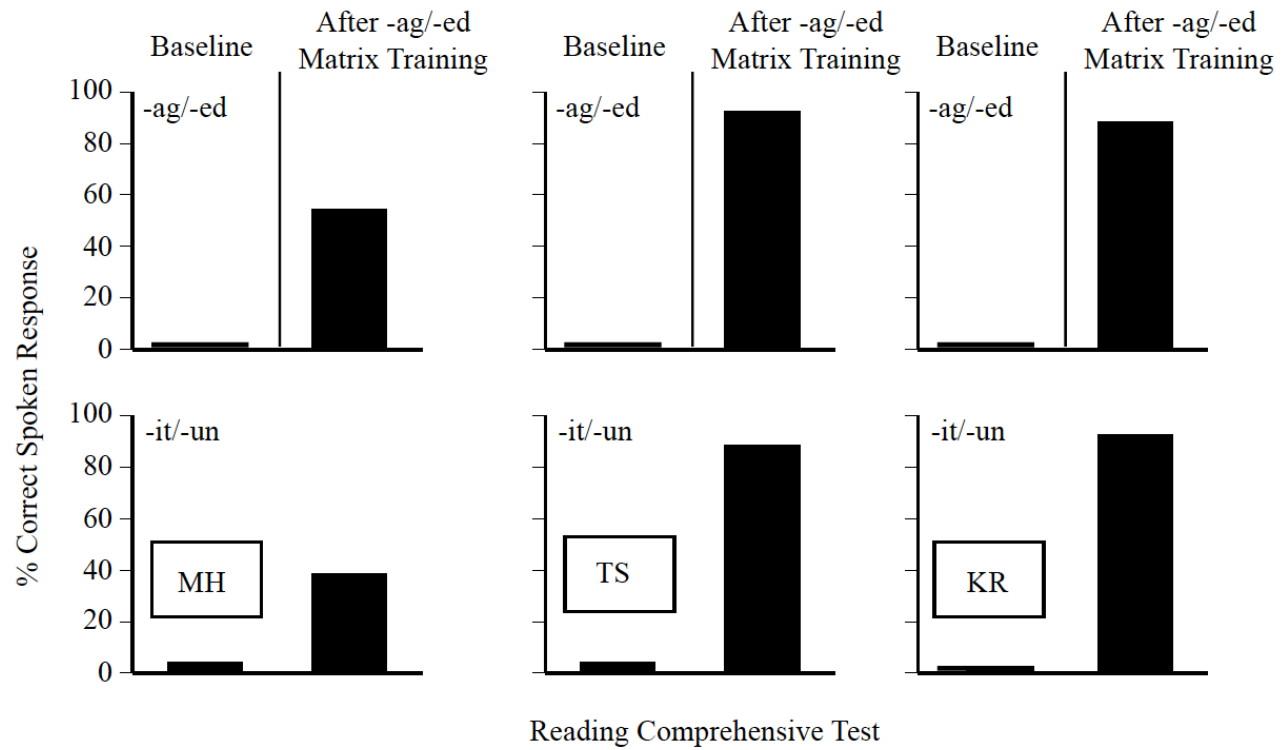


Figure 5. The percentage of correct responses made by MH (left), TS (middle), and KR (right) during the Reading Comprehension Test are shown. The vertical black line represents -ag/-ed Matrix Training.

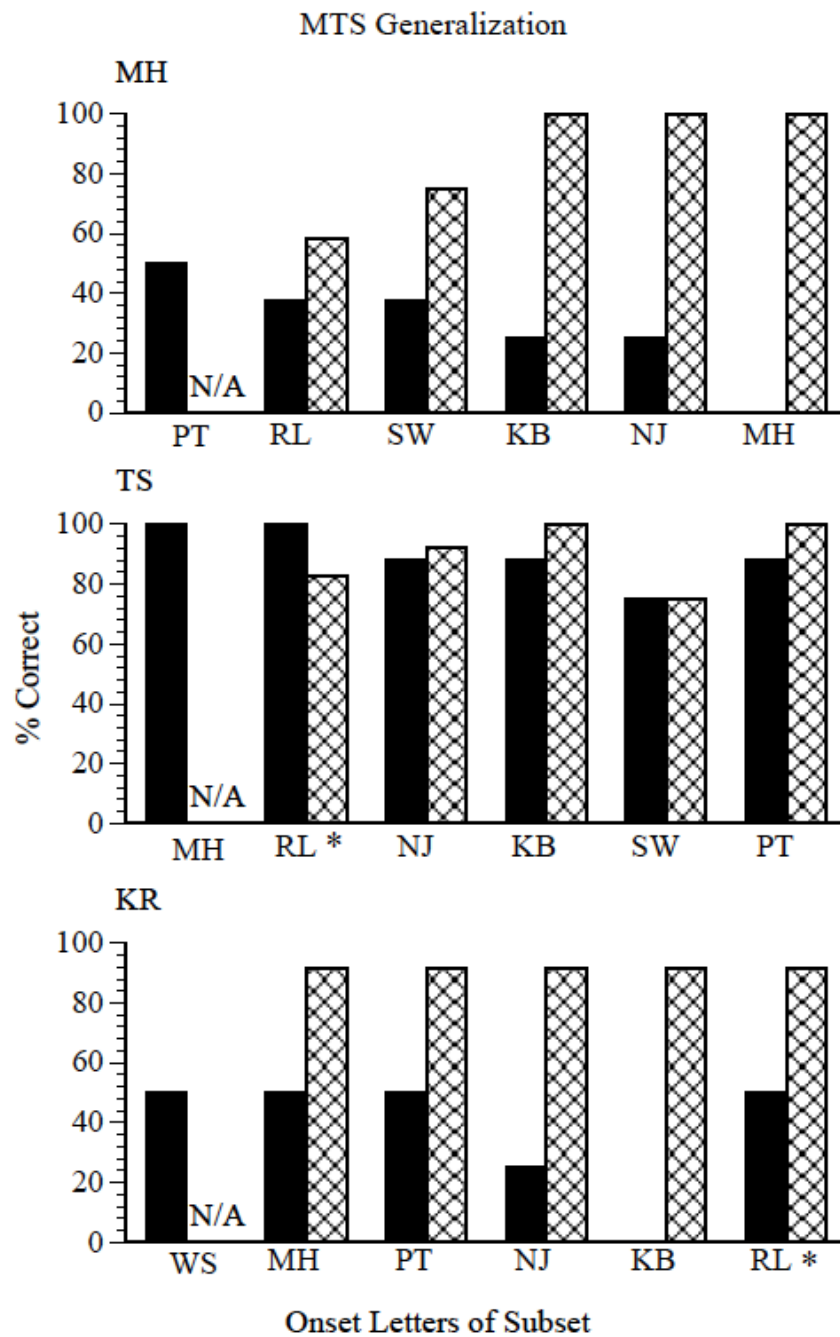


Figure 6. The percentage of correct responses made by MH (top), TS (middle), and KR (bottom) on the baseline MTS Comprehensive Test (black columns) and the MTS Generalization Tests (crossed columns) given during Matrix Training for each subset in the -ag/-ed matrix. Subsets are shown in order of training from left to right on the x-axis. *Accuracy on the RL subset for TS and KR are based on rimes.

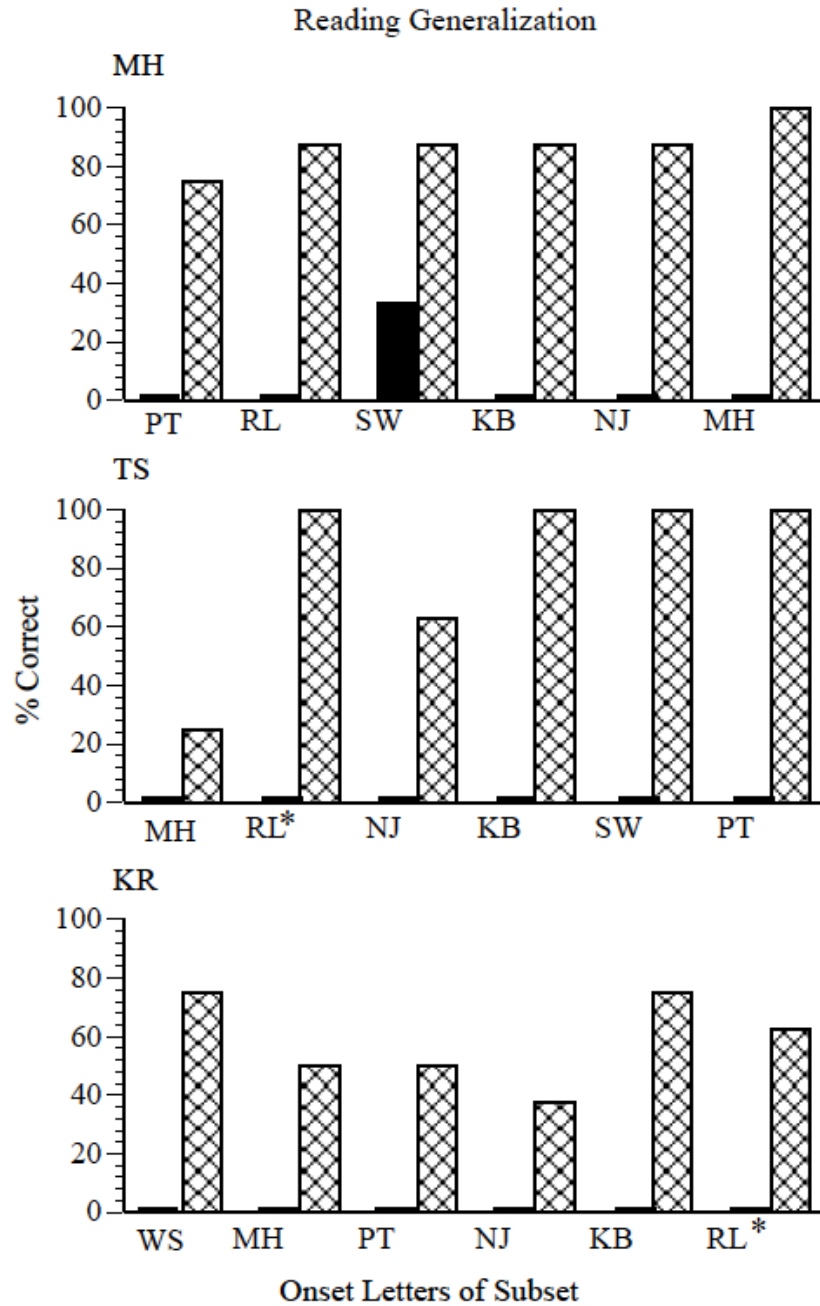


Figure 7. The percentage of correct responses made by MH (top), TS (middle), and KR (bottom) on the baseline Reading Comprehension Test (black columns) and the Reading Generalization Tests (crossed columns) given during Matrix Training for each subset in the -ag/-ed matrix. Subsets are shown in order of training from left to right on the x-axis. *Accuracy on the RL subset for TS and KR are based on rimes.

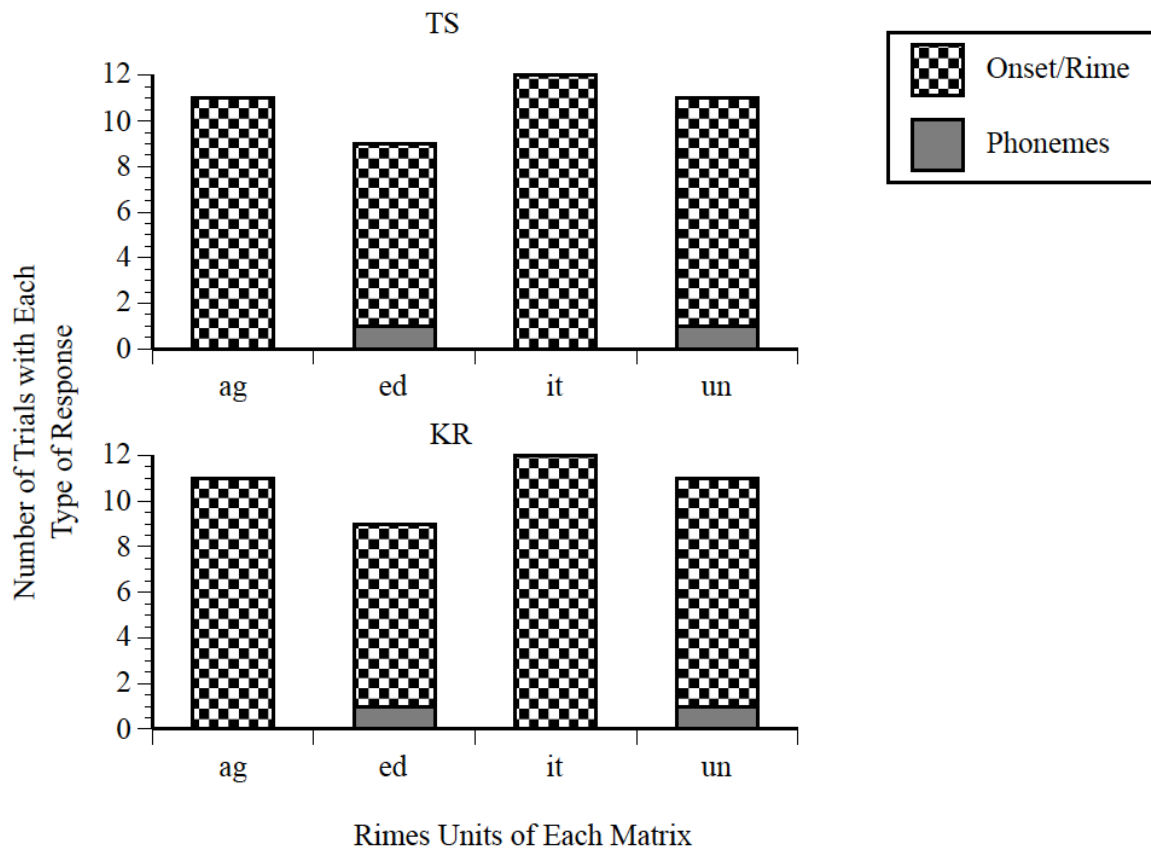


Figure 8. The number of whole word responses (out of 12 opportunities) that were preceded by subsyllable sounds (y-axis) during the Reading Comprehension Test given following mastery of the –ag/-ed matrix. TS is shown on the top panel, and KR on the bottom. The rime units of each matrix are presented across the x-axis. Each column shows the number of whole word responses that were preceded by phonemes (grey) and onset/rime sounds (checkered). This figure does not reflect accuracy of responses.

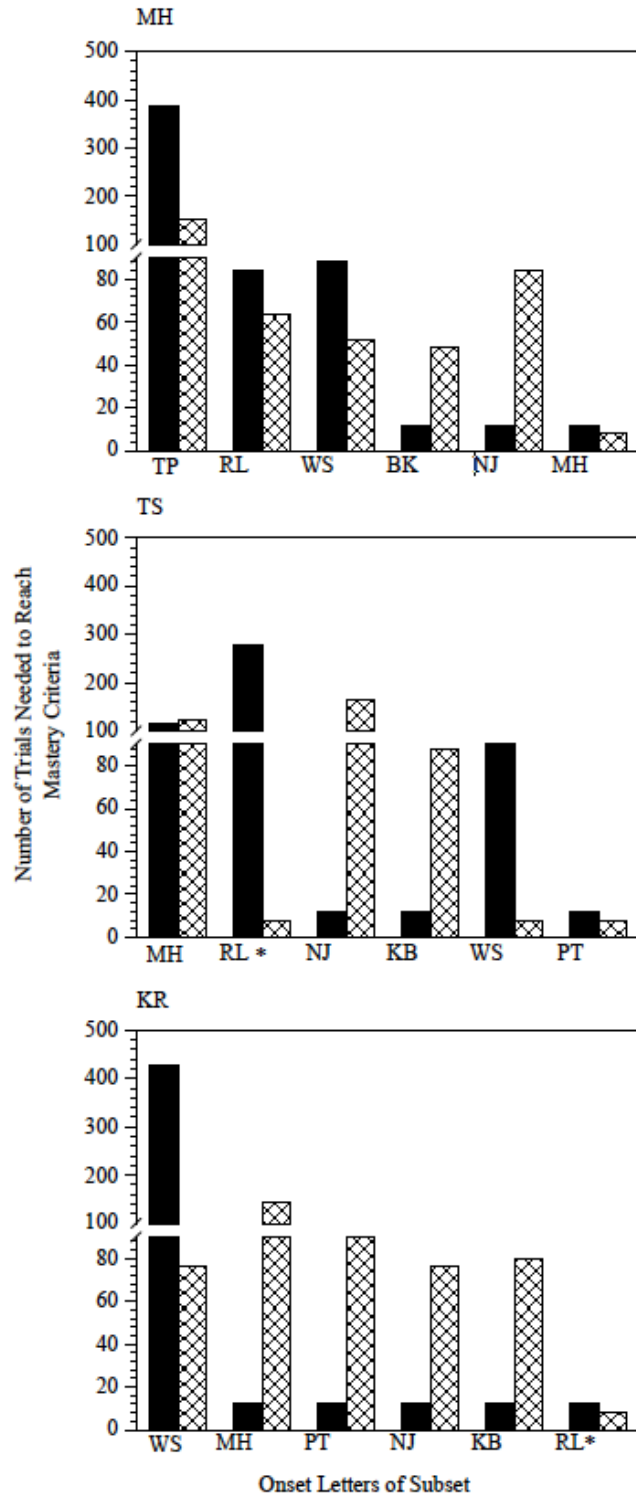


Figure 9. The number of trials (y-axis) needed for MH (top), TS (middle), and KR (bottom) to reach mastery criteria on MTS (black columns) and Reading (crossed columns) for each subset (x-axis). Subsets are indicated by their onsets and are shown in order of training from left to right. *Accuracy on the generalization tests for the RL subset for TS and KR were based on rimes only.

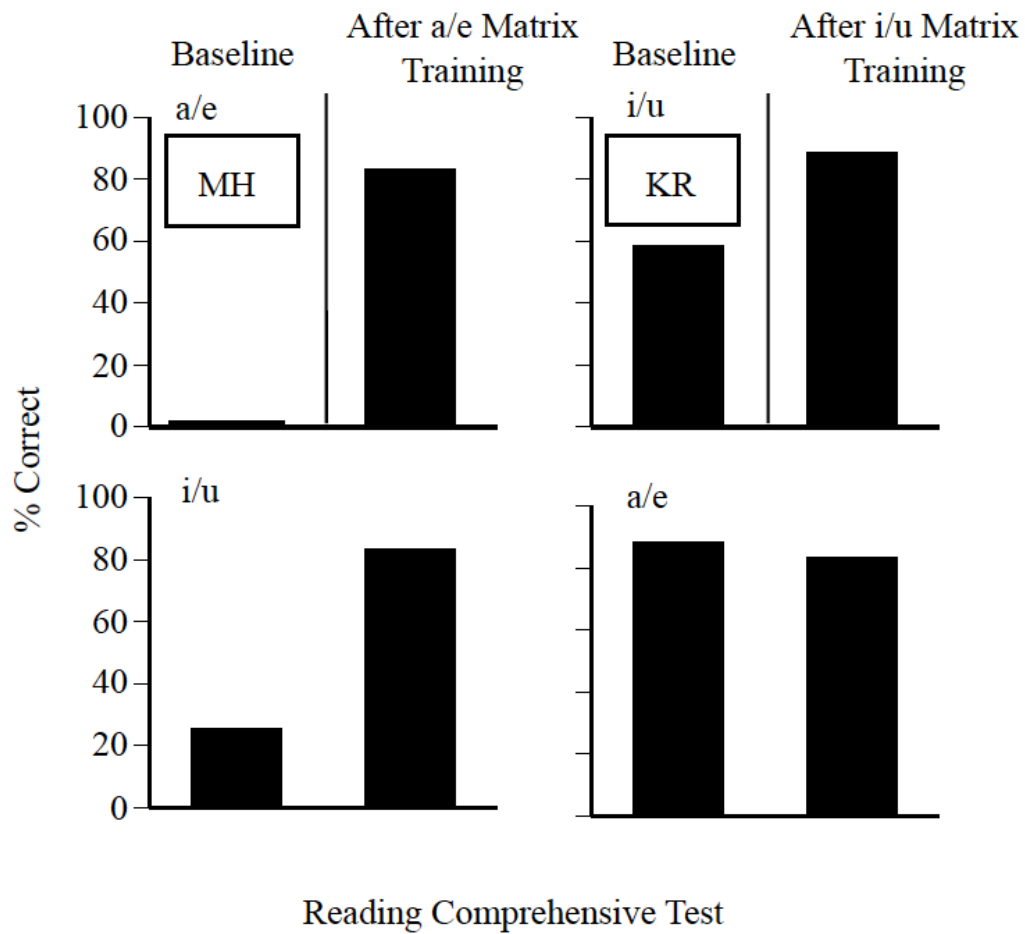


Figure 10. Accuracy on the Reading Comprehension Test before and after training for MH (left) and KR (right). Matrices are indicated by the vowels at the top left of each panel. Training is indicated by a black vertical line.

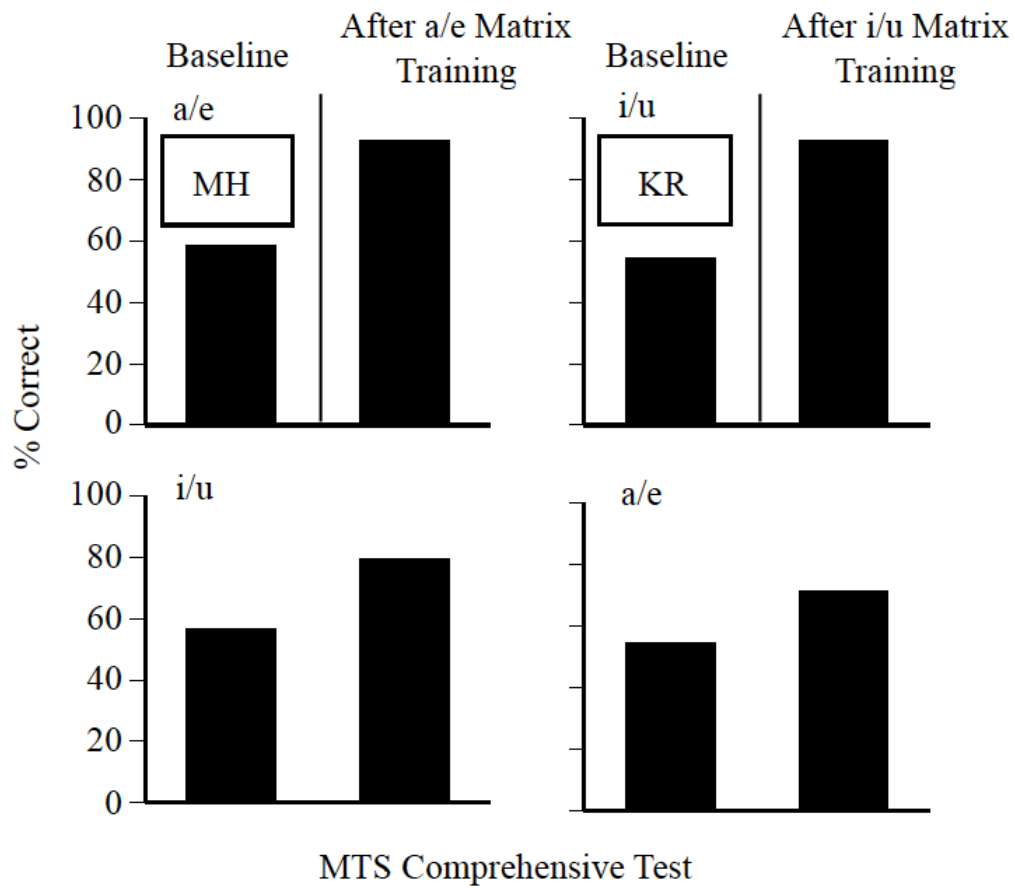


Figure 11. Accuracy on the MTS Comprehensive Test before and after training for MH (left) and KR (right). Matrices are indicated by the vowels at the top left of each panel. Training is indicated by a black vertical line.

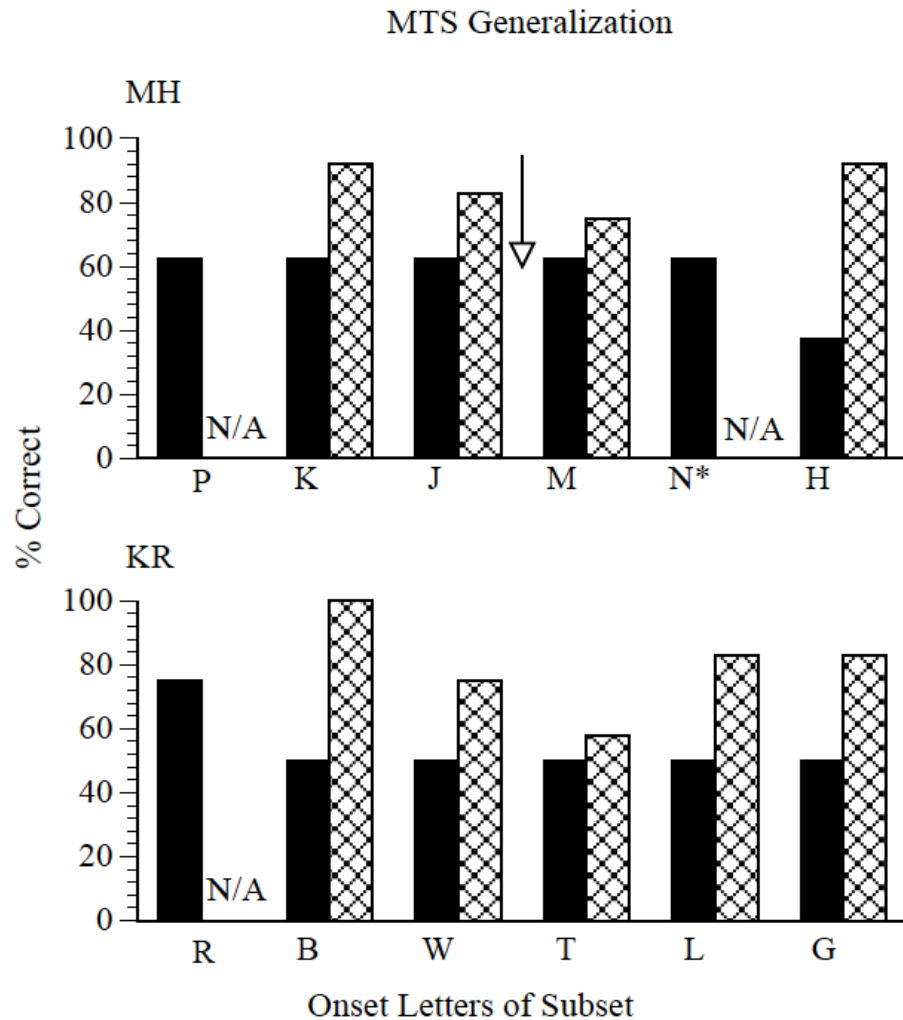


Figure 12. Accuracy on the baseline MTS Comprehensive Test (black columns) and the MTS Generalization Test (crossed columns) for each subset, indicated by the onset across the x-axis, for MH (top) and KR (bottom). The arrow in MH's graph represents when the Reading Probe was given. *There was not an MTS Generalization Test for the N Subset.

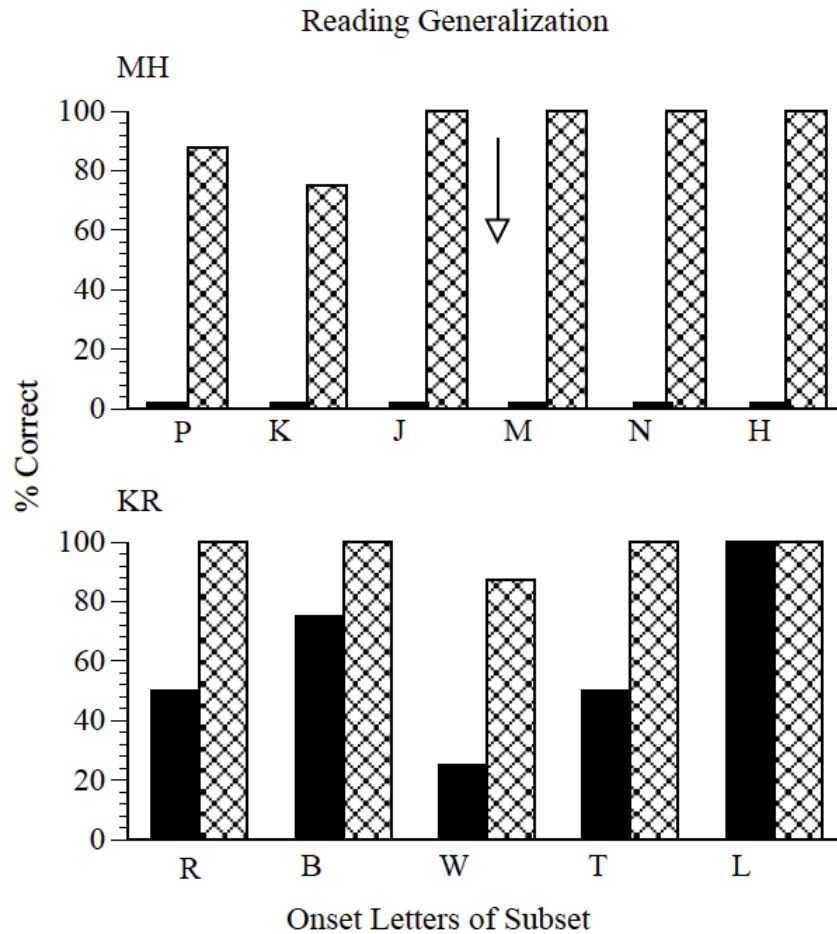


Figure 13. The percentage of correct responses made by MH (top) and KR (bottom) on the baseline Reading Comprehension Test (black columns) and the Reading Generalization Tests (crossed columns) given during Matrix Training for each subset trained in Study 2. Subsets are shown in order of training from left to right on the x-axis. The arrow in MH's graph represents when the Reading Probe was given (see Appendix).
 *Due to time constraints, KR did not receive MTS training on the G subset, therefore his Reading Generalization Test score for that subset is not shown.

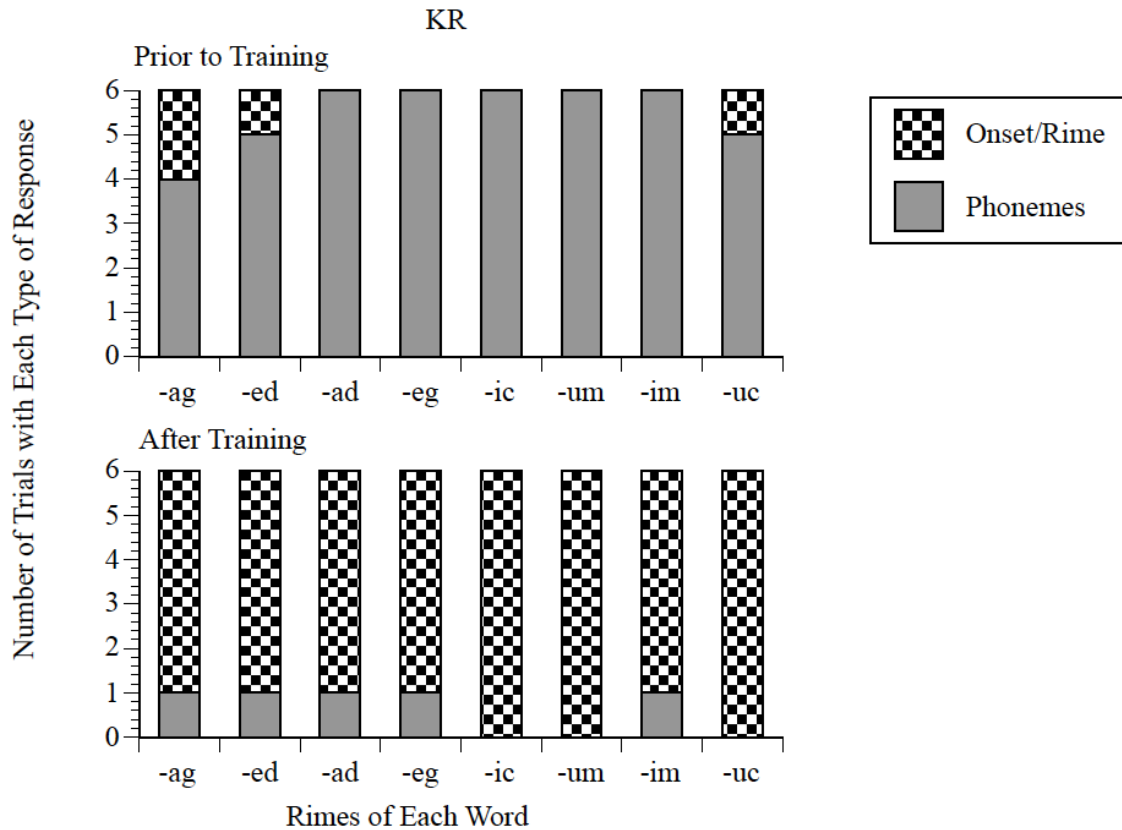


Figure 14. The number of KR's whole word responses (out of six opportunities) that were preceded by subsyllable sounds during the Reading Comprehension Test given before (top) and after (bottom) training. The rime units of each matrix are presented across the x-axis. For each rime there is a column showing the number of responses to words ending in that rime that were preceded by phonemes (grey), onset/rime sounds (checkered). This figure does not reflect accuracy of responses.

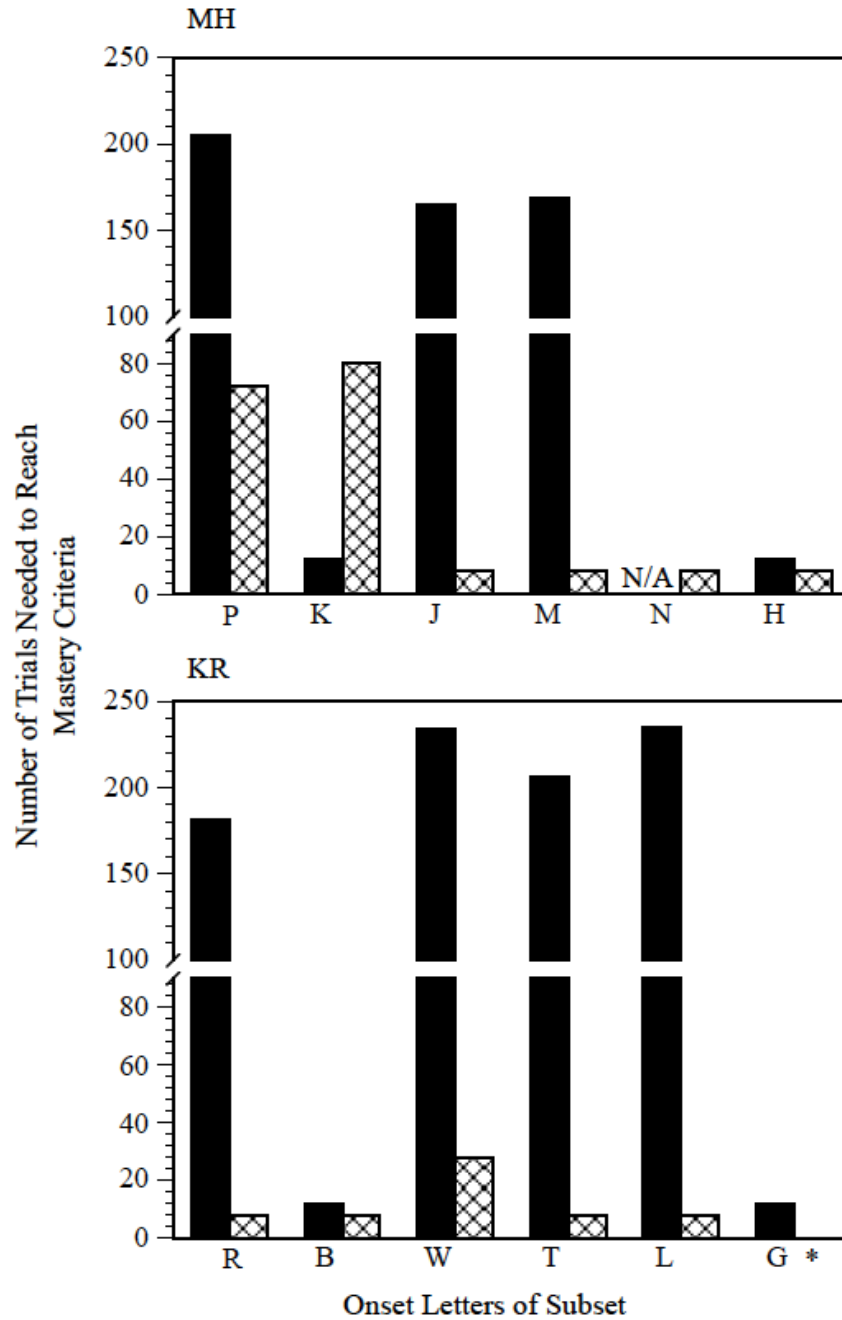


Figure 15. The number of trials (y-axis) needed for MH (top) and KR (bottom) to reach mastery criteria on MTS (black columns) and Reading (crossed columns) for each subset (x-axis). Subsets are indicated by their onsets and are shown in order of training from left to right. *G subset was not trained for reading due to time constraints.

Appendix A

Study 1

MTS Training

Prompted Trials Prompted trials (Figure 3) were identical to other trials except the printed word corresponding to the spoken word was presented in the center of the screen while the choices were presented at the corners.

Full MTS Training All participants received Full MTS Training for the first subset trained, which varied across participants according to accuracy on the baseline MTS Comprehensive Test. Each step described below began with a prompted trial for each word presented in the step. Criterion to move out of prompted trials depended on the number of words being trained in each step. If there were two words being trained in a step, the participant had to respond correctly on two consecutive prompted trials to move on (one prompt for each word). If there three words being trained, the participant had to respond correctly on three consecutive prompted trials, and so on.

Step 1: Teach a word pair differing in onsets only. Two words, differing only by the onset (e.g. rag, lag), were trained in a two-choice format. Following correct responses on two consecutive prompted trials, participants had to respond at 100% accuracy on eight consecutive trials to move on to Step 2.

Step 2: Teach a word pair differing in rimes only. In this step, the words had the same onset with different rimes (e.g. rag, red). The criteria for moving from prompted to unprompted trials, as well as moving into the combination task, were identical to Step 1.

Step 3: Teach a combination of 3 words. This three-choice task trained two words with the same onset and one word with a different onset (e.g. lag, led, red). Two words had the same onset and two words had the same rime. After responding correctly on three consecutive prompted trials, participants had to respond correctly on 11 of 12 consecutive trials to move on to Step 4.

Steps 4, 5, and 6: Teach remaining contrasts. The first three steps were repeated in the same order, with the same criteria using the remaining two and three word combinations from the subset.

Step 7: Teach all words together. All the words in the subset were presented in a four-choice task. Participants had to respond correctly on four consecutive prompted trials, then had to respond correctly on 11 of 12 consecutive trials correct for two consecutive days. The second day showed maintenance of the words in the absence of prompts for at least 24 hours.

Expedited MTS Training for TS. Expedited MTS Training was used with TS only. His previous experience with MTS led to high scores on the baseline MTS Comprehensive Test, so we focused mainly on his generalization from MTS to reading. We used Expedited MTS Training so TS would have a review of the MTS task immediately prior to the Reading Generalization Test for each subset. TS' training combined Steps 1-3 and Steps 4-6 of full MTS Training by using two trial types in each step. We will refer to this as a mixed-trial format.

Step 1 & 4: Mixed-trial. For onset training, one trial type had one pair of words containing the same rime (e.g. rag, lag) and the other trial type had the remaining pair

(e.g. red, led). After responding correctly on four consecutive prompted trials, TS had to respond correctly on eight consecutive trials to move on to the next step.

Step 2 & 5: Mixed-trial. This step was identical to the previous except it focused on the rime unit: one trial type had a pair of words containing the same onset (e.g. lag, led) and the other trial type had the remaining pair (e.g. rag, red). Criteria were the same as in the previous step.

Step 3 & 6: Mixed-trial. For this three-choice task, the four words of the subset were organized into two groups. Each group had two words with the same onset and two words with the same rime (e.g. rag, red, led & lag, led, rag), creating two trial types. TS had to respond correctly on six consecutive prompted trials before moving into the unprompted trials, which required 11 correct responses out of 12 consecutive trials in order to move on to Step 7.

Step 7: Teach all words together. The four-choice task for expedited training was identical to that used in Full MTS training.

Appendix B

Study 1

Reading Training

Verbal Contrast Reading Training On the first trial of the first verbal contrast task, the experimenter said : “You will see a word on the screen, I will give you two choices and I want you to tell me which one it is.” These instructions were not repeated on following trials. On each trial, if a participant said a word that was not one of the two provided by the experimenter, the response was recorded and the experimenter repeated the two choices to the participant until s/he said one.

If the participant did not make a response within 5s, the experimenter prompted by repeating the instructions and waited another 5s before providing the correct word. The participant was then asked to repeat the correct word. Following an incorrect response, the experimenter provided the same corrective feedback. Correct answers always produced verbal praise and delivery of the reinforcer. If criterion was not met on a step after 32 trials, the participant moved back a step.

Step 1: Verbal onset contrast. During onset contrast, the experimenter said the correct word and another word from the subset that shared the same rime that differed only by the onset. For example, if rag were displayed on the screen, the experimenter said “rag or lag?” The participant had to choose the correct word on two consecutive blocks of four (8 trials), each block containing all the words in the subset, in order to move to the next step.

Step 2: Verbal rime contrast. The next task was identical to the first, except the experimenter said the correct word and a distractor in the subset with the same onset, but different rime (e.g. “rag or red?”).

Step 3: Verbal onset & rime contrasts. The following task was identical to the first two, except that the blocks of four alternated between onset contrast and rime contrast. In the first block of four words, the experimenter used the onset contrast procedure, then used the rime contrast procedure on the following block of four, and so on.

Unprompted Reading Training. The final task in reading training required participants to produce the word printed on the screen with no verbal prompts from the experimenter. Criterion to move on was the same as in the verbal contrast tasks, except that criterion had to be reached on two days. The second was a day with no verbal prompting. That is, the participant could not have a session with prompts on the same day as Unprompted Reading Training.

Appendix C

Study 2

MTS Training

Step 1: Teach a word pair differing in rimes only. A pair of words from the first subset containing rimes with no overlap (e.g. red, rag) was trained in a two-choice task. Following correct responses on two consecutive prompted trials, participants had to respond at 100% accuracy on eight consecutive trials to move on to Step 2.

Step 2: Teach a word pair differing in codas only. This step was identical to the first, except the words only differed only by coda (e.g. rag, rad). Criteria for moving out of prompted trials, as well as moving into Step 3, were identical to Step 1.

Step 3: Teach a word pair differing in vowels only. This step was identical to the first and second, except the words differed only in the vowel (e.g. rad, red). Criteria to move out of prompted trials, as well as moving to Step 4, were identical to the previous steps.

Step 4: Teach a combination of 3 words. This three-choice task trained two words with the same vowel and two with the same coda, requiring a comparison of two vowels and a comparison of two codas (e.g. rag, rad, red). Following correct responses on three consecutive prompted trials, participants had to respond correctly on 11 out of 12 consecutive trials in order to move on to Step 5.

Steps 5-8: Teach remaining contrasts. Once the participants completed the first four steps, those steps were repeated in the same order with the same criteria using the remaining two and three word combinations from the subset.

Step 9: Teach all words together. Following the completion of Step 8, participants completed a four-choice task presenting all the words in the subset before moving on to the Reading Generalization Test. Criterion on the four-choice task was 11 out of 12 consecutive trials correct for two consecutive days, the last of which had no prompted trials at the beginning. The last day served to show participants could maintain accuracy without prompted trials.

Appendix D

Study 2

Reading Training

Verbal Comparison Reading Training. These methods were similar to those used in Study 1 with steps added because of the higher amount of overlap between words in the new matrices. On each trial, a word was presented on the screen, and the experimenter said two words: the word on the computer screen and an incorrect word. Two pairs of words from the same subset were presented across alternating trials. Participants move to the next step when they respond correctly on two consecutive blocks of four, each block containing all the words in the subset.

Step 1: Rime contrast. The words shared no overlap in the rime. For example, if pag was displayed on the screen, the experimenter said “pag or ped?” One trial type presented words ending in -ag/-ed and the other presented words ending in -eg/-ag. Participants had to respond correctly on eight consecutive trials to move on to Step 2.

Step 2: Coda contrast. The words had the same vowels but different codas, e.g. saying “pad or pag?” given the printed word pag. Criterion to move on to Step 3 was identical to Step 1.

Step 3: Vowel contrast. The words had the same coda, but different vowels, e.g. saying “pad or ped?” in the example provided above. Criterion to move on was identical to the previous steps.

Step 4: Rime, coda, and vowel contrasts. Step 4 was identical to the first three, except the experimenter used the rime contrast procedure on the first block of four words, the coda contrast procedure on the following block of four, the vowel contrast procedure,

and so on. The sequence repeated until the participant responded correctly on 11 out of 12 consecutive trials to move on to Step 5.

Step 5: Unprompted Reading Training. In the last step, the experimenter offered no spoken words for the participant to choose from. Criterion on the four-choice task was 11 out of 12 consecutive trials correct for two consecutive days, the last of which had no prompted trials at the beginning. The last day served to show participants could maintain accuracy without prompted trials.

Error Contrast Reading Training. This is a condensed version of the Verbal Comparison Reading Training, used only when there are two or less errors made on the Reading Generalization Test. The type(s) of error(s) made dictate(s) the contrast that was used: rime, vowel, or coda. The correct word and the incorrect response provided by the participant are compared to each other. In order to move on to Unprompted Reading Training, participants must respond correctly on one block of four, each block containing both words twice.

If a participant made more than one type of error on the Reading Generalization Test for a subset, he completed training for each error made. In the case that, on the Reading Generalization Test, a response was given that is not one of the words in the current subset, the error made still dictated what type of comparison was used. For example, saying “bed” for the printed word pag was considered a rime error and the words to be compared would be pag and ped.

Rime error comparison. When a rime error was made (e.g. saying pag for ped), the pair of words with contrasting rimes containing the word missed was presented in a verbal two-choice format; the experimenter said the word missed and the word in the

subset with the contrasting rime. For example, if pag were displayed on the screen, the experimenter said “pag or ped?”

Vowel error comparison. Vowel comparison training was identical to rime comparison training, except that it only applies when the response made only differs from the correct response in the vowel (e.g. saying pag for peg).

Coda error comparison. Coda comparison training is identical to rime comparison training, except that it only applies when the response made only differs from the correct response in the coda (e.g. saying pag for pad).

Unprompted Reading Training. This was identical to Step 5 of Verbal Comparison Reading Training.

Appendix E

Reading Scoring Procedures

Scoring procedures were the same across both studies,. The experimenter used a premade scoresheet tailored to the specific reading task for that session (see Figure 20). On the scoresheet were spaces for the date, the experimenter's name, and the participant's ID. There was also a table with columns containing (from left to right): trial number, word presented on the screen, the participant's response (recorded during session with pencil), accuracy of whole-word response (1 = correct, 0 = incorrect), one column for accuracy on each phoneme (CVC), and a column for notes. Scoresheets for the Verbal Comparison Training in Study 2 had an additional column indicating what two words to say to the participant on each trial.

See Figure 20 for an example of the scoring procedure for possible responses given by participants. The experimenter recorded all spoken responses, relevant to the task, during the session. An example of a response that is not relevant to the task is, "My dog likes to wag its tail sometimes" or "When will we be done?" Responses recorded include subsyllable sounds (e.g. /b/, /ag/, /ba/) and whole-word responses. All responses were recorded, even if there is more than one on a particular trial.

The first whole-word response given by the participant is circled by the experimenter and scored. Following responses are recorded, but do not count towards accuracy. Each subsyllable sound is recorded with forward slashes surrounding the corresponding letter(s) (e.g. /b/, /at/). If a participant says individual letters, the experimenter records each letter separated by a dash (e.g. b-a-t). If a participant does not

respond within the given amount of time, a dash is recorded in the Participant's Response column, and "No Response" is written in the Notes column for that trial.

Experimenters were trained to score the reading tasks according to the following procedures. After reading the protocol and discussing any questions, the experimenters conducted a role-playing exercise. A mock session was conducted, during which Experimenter A pretended to be a participant, and Experimenter B had to score each response. Afterwards, Experimenter A looked over the completed scoresheet and offered feedback on Experimenter A's performance. They then switched. This continued until both were fluent in the procedures and there were no disagreements.

Nature of error	Word	P's Response	Score				Notes
			CVC	C	V	C	
Add syllable	Met	mee-um-um	0	1	0	0	
	pan	pannies	1	1	1	1	
	pad	pero-eh-eh	0	1	0	0	
Single consonant vs. cluster	pug	plug	0	0	1	1	
	win	wing	0	1	1	0	
Long vowel, drop coda	mid	my	0	1	0	0	
Long vowel	pip	pipe	0	1	0	1	
Common substitutions	rug	wug	1	1	1	1	wr
	dad	bad	1	1	1	1	bd
Individual phonemes (letter sounds)	bug	/b/ /u/ /g/	0	0	0	0	
Says each letter individually	bat	b-a-t	0	0	0	0	
not understandable; bad articulation	mag	/m/	0	-	-	-	articulation
Self Correction// Multiple Response	mag	med mag mag	0	1	0	0	

Figure 16. A scoresheet, with added column for "Nature of Error," used to show the possible spoken responses a participant can give and how the experimenter should record them during the session.